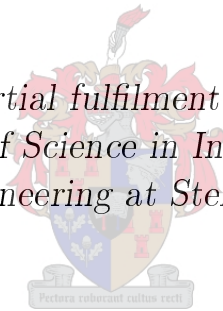


A RISK-RELIABILITY COMPARISON OF TRACK SECTIONS IN THE PASSENGER RAILWAY INDUSTRY

by

M.D. Mc Naught

*Thesis presented in partial fulfilment of the requirements for
the degree of Master of Science in Industrial Engineering in
the Faculty of Engineering at Stellenbosch University*



Department of Industrial Engineering,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof C.J. Fourie & Mr K.H. Von Leipzig

December 2015

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not, previously, in its entirety or in part, submitted it at any university for a degree.

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature:

Date:

Copyright © 2015 Stellenbosch University
All rights reserved

Abstract

In chaotic maintenance environments, executing planned maintenance becomes difficult because the need for immediate corrective action escalates. Reverting back from a poor system state to a stable and well maintained one is a challenge. Railway track environments are prone to system degradation and poor maintenance. They are, therefore, in need of analytical tools to ‘get on track’ with maintenance.

A risk-based method which grades track corridors between train stations according to their level of risk was developed. To achieve this, both the likelihood and the severity components of risk were considered. The likelihood component of risk in the track environment is the reliability of track. Reliability was calculated by quantifying track failure modes first and then analysing the characteristics of failures for each track corridor. Probabilistic models were generated from repairable systems reliability theory from which reliability predictions were made. The severity component of risk is the average delay historically experienced by each track corridor. A risk matrix was developed which brings together likelihood and severity components of risk for each track corridor. Maintenance prioritisation is possible from the risk rankings created by the matrix. The risk rankings for five track corridors were validated when a condition-based track maintenance tool, TQI, was in agreement.

Uittreksel

In die chaotiese instandhoudingsomgewing is dit moeilik om beplande onderhoud uit te voer, aangesien die nodigheid vir onmiddellike regstellende aksie toeneem. Om terug te keer van 'n swak stelsel na 'n stabiele en goed instandgehoue stelsel is 'n uitdaging. Spoorweg-omgewings is geneig tot stelsel-agteruitgang en swak instandhouding. Daar is dus 'n behoefte aan analitiese metodes om weer op die regte skedule te begin volg met instandhouding.

'n Risiko-gebaseerde metode wat spoorweë tussen stasies gradeer volgens hulle risiko is ontwikkel. Om dit reg te kry, is beide die waarskynlikheid en die graad van erns van risiko's in ag geneem. Die waarskynlikheid van risiko in die spoor-omgewing is die betroubaarheid van die spoor. Betroubaarheid word bereken deur eerstens die hoeveelheid spoorfalings te bepaal en dan die kenmerke van die falings vir elke spoorweg te analiseer. Waarskynlikheidsmodelle is opgestel van herstelbare stelsels betroubaarheidsteorie van waar betroubaarheidsvoorspellings gemaak is. Die graad van erns van risiko's is die gemiddelde vertraging wat histories deur elke spoorweg ondervind is. 'n Risiko-matriks wat die waarskynlikheid en graad van erns van risiko's kombineer is vir elke spoorweg ontwikkel. Instandhoudings-prioritisering word moontlik gemaak deur die risiko-graderings volgens die matriks. Die risiko-graderings vir vyf-spoor spoorweë is bevestig toe 'n voorwaarde-gebaseerde instandhoudingsprogram, TQI, eenparigheid bereik het.

Acknowledgements

I would like to express my gratitude to the following individuals who have given selflessly for my gain during the course of my project:

To Professor Fourie, for his encouragement, wisdom and guidance. He treated me like a son in every work endeavour I undertook whilst being a member of his office. To Lisa, for her love and appraisal, which gave me the confidence to persevere when I thought all hope was lost. I thank her mostly for her grace, for the attention I gave my project that she so rightfully deserved. To my dad, for always providing and contributing to my success. I had a full stomach during late nights because of you. To the rest of my family, for giving me the space I needed to focus, in the midst of a distracting year. To 'Bossie' and Mackson, for providing data, which was critical for the realism of my case study. To Rasheed, for giving me hope that quality individuals exist at the bottom. Importantly, thanks to Pieter, for his teasing and concern over my work. The robustness of my project is thanks to you. To God, who is far beyond an individual, who gave me strength.

Contents

Declaration	i
Abstract	ii
Uittreksel	iii
Acknowledgements	iv
Contents	v
List of Figures	viii
List of Tables	xi
Nomenclature	xii
1 Introduction	1
1.1 Background	1
1.2 Problem statement	2
1.3 Research objectives	2
1.4 Research design and methodology	3
2 Literature Review	5
2.1 Infrastructure	5
2.2 Reliability of perway components	10
2.3 Track quality index	21
2.4 Effects and criticality	26
2.5 Root cause analysis	29
2.6 Track service life	31
2.7 Data cleaning	32
2.8 Repairable and non-repairable system reliability	34
2.9 Two important reliability theory functions	35
2.10 Statistical analysis of system failure data	36
2.11 Trend testing	37
2.12 Point process models	40

2.13	Parameter estimation	42
2.14	Goodness of fit	45
2.15	Risk analysis	47
2.16	Summary of solution methods	51
3	Case Study: PRASA Infrastructure Maintenance Priority Using Risk and Reliability	53
3.1	Maintenance investigation	54
3.2	Track corridor reliability model	57
3.3	Failure mode identification and criticality analysis	59
3.4	Root cause analysis	59
3.5	Risk matrix	61
4	Data Analysis	66
4.1	Data sets	66
4.2	Reliability methods applied	68
4.3	Pareto analysis of perway failure mode severity	77
4.4	Risk matrix applied	79
5	Results and Validation	82
5.1	Reliability analysis of multiple perway corridors	82
5.2	Risk analysis of multiple perway corridors	85
5.3	Validation of risk matrix	88
5.4	Sensitivity analysis of selected risk models	90
6	Conclusions and Recommendations	91
6.1	Back to the research problem	91
6.2	Theoretical contribution and practical application of methods	92
6.3	Conclusions about limiting factors encountered	93
6.4	Future work	94
	Appendices	96
A	FMEA Tables	97
B	Fault Trees for Top Ten Failure Modes	100
C	Cumulative Failure Against Time for Five Perway Corridors	106
D	Linear Estimation Method for Five Perway Corridors	109
E	Confidence Bounds on ROCOF for Five Perway Corridors	114
F	Urban Transport 2015 Conference Article	119
G	South African Journal of Industrial Engineering Article	132

CONTENTS

vii

List of References

150

List of Figures

1.1	Research methodology for the development and validation of a reliability-based risk perway section maintenance prioritisation model.	4
2.1	Simplified rail profile for a section of perway.	10
2.2	Risk of a defect type developing into a rail break for progressive age of rail (Kumar <i>et al.</i> , 2010).	11
2.3	Reliability block diagram for a 200 meter section of perway.	15
2.4	Methodology to determine train derailment probability (Mohammadzadeh and Ghahremani, 2012).	17
2.5	k-out-of-n reliability analysis for $k = 2$ and $k = 3$ systems (Zhao <i>et al.</i> , 2007).	18
2.6	Rail rollover failure mode (Iwnicki, 2006).	19
2.7	Top view of train loaded rail with rail clip supports.	19
2.8	Moment diagram of rail section with train forces and clip forces (Iwnicki, 2006).	20
2.9	Plasser IM2000 geometry measuring rail car (Zaayman, 2013).	22
2.10	Averaged estimate for change in TQI between maintenance intervals (Sadeghi and Askarinejad, 2011).	24
2.11	Relationship between TQI and rail derailment probability due to track irregularity.	26
2.12	Methodology for the construction of FMECA for railway perway.	28
2.13	Methodology for the construction of fault tree analysis for railway perway.	30
2.14	Framework for data cleaning (Van den Broeck <i>et al.</i> , 2005).	34
2.15	Counting process for a number of system failures.	36
2.16	Statistical analysis decision tool for a point process (Ascher and Feingold, 1984).	37
2.17	Standard normal distribution used in Laplace test.	38
2.18	Risk matrix varieties (Markowski and Mannan, 2008).	48
2.19	Risk in the business context (Ng <i>et al.</i> , 2003).	50
3.1	Reduced map of the Cape Town perway network.	55
3.2	Database entry illustrating asset management levels for perway.	56
3.3	Reliability model of train network.	57

3.4	Reliability model of train network.	58
3.5	Fault tree for rail break failure mode.	62
3.6	Risk matrix model for perway maintenance prioritisation.	64
4.1	Line map extract of the Metrorail perway network.	67
4.2	Combined work order and events database extract between Fish Hoek and Kalk Bay.	68
4.3	Number of perway failures for Metrorail Cape Town.	69
4.4	Decision support diagram for the location identification of perway failure events.	71
4.5	Cumulative failure against time for the Simonstown - Glencairn perway corridor.	72
4.6	Log-linear and power law model fit to failure data from the Simon- stown - Glencairn perway corridor.	73
4.7	Estimated ROCOF for a) log linear and b) power law models. . . .	74
4.8	ROCOF with 95% upper and lower confidence bounds for a) log linear and b) power law models.	75
4.9	Reliability of power law and log linear models for the perway failures of the Simonstown - Glencairn corridor.	76
4.10	Theoretical degradation of system reliability with asset reliability for Simonstown - Glencairn perway corridor.	77
4.11	Pareto diagram of train delay caused by each perway failure mode. .	78
4.12	Risk matrix for the Simonstown - Glencairn perway corridor. . . .	81
5.1	Time intervals over which trends were established for five perway track corridors.	83
5.2	Reliability against time for five perway corridors in the Metrorail Cape Town network.	85
5.3	Risk matrix for the comparison of five perway corridors in the Metrorail Cape Town network.	86
B.1	Fault tree for rail clip failure mode.	100
B.2	Fault tree for sand on tracks failure mode.	101
B.3	Fault tree for track geometry twist failure mode.	101
B.4	Fault tree for block joint failure mode.	102
B.5	Fault tree for railway crossing failure mode.	102
B.6	Fault tree for track geometry horizontal alignment failure mode. . .	103
B.7	Fault tree for rail crown failure mode.	104
B.8	Fault tree for dirty ballast failure mode.	105
C.1	Cumulative failure against time for the Kalk Bay - Fish Hoek per- way corridor.	106
C.2	Cumulative failure against time for the Philippi - Nyanga perway corridor.	107

C.3	Cumulative failure against time for the Maitland - Ndabeni perway corridor.	107
C.4	Cumulative failure against time for the Southfield - Heathfield perway corridor.	108
D.1	Estimated ROCOF for Kalk Bay a) log linear and b) power law models	110
D.2	Estimated ROCOF for Philippi a) log linear and b) power law models	111
D.3	Estimated ROCOF for Maitland a) log linear and b) power law models	112
D.4	Estimated ROCOF for Southfield a) log linear and b) power law models	113
E.1	ROCOF of the Kalk Bay perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models. . . .	115
E.2	ROCOF of the Philippi perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models. . . .	116
E.3	ROCOF of the Maitland perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models. . . .	117
E.4	ROCOF of the Southfield perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models.	118

List of Tables

2.1	Key contributors leading to track geometric defects (Sadeghi and Askarinejad, 2009).	23
2.2	Rail service life (Mundrey, 2010).	32
3.1	RPN table for railway perway failure modes (Chin <i>et al.</i> , 2009), (Xu <i>et al.</i> , 2002).	60
3.2	Identified preventative maintenance actions for ten specified failure modes.	63
4.1	Failure modes and their average delay experienced on the Simonstown - Glencairn perway corridor.	80
5.1	Trend testing and curve fitting for five perway corridors.	84
5.2	TQI values and risk metrics for five perway corridors in the Metro-rail Cape Town network.	89
5.3	Sensitivity for risk and reliability comparison of log linear and power law models.	90
A.1	FMEA occurrence evaluation (Chin <i>et al.</i> , 2009), (Xu <i>et al.</i> , 2002).	97
A.2	FMEA severity evaluation (Chin <i>et al.</i> , 2009), (Xu <i>et al.</i> , 2002).	98
A.3	FMEA detection evaluation (Chin <i>et al.</i> , 2009), (Xu <i>et al.</i> , 2002).	99
C.1	Sensitivity for risk and reliability comparison of log linear and power law models.	108

Nomenclature

Variables

c_i	Average measurement of the j_{th} track irregularity parameter
$f_x(\vartheta)$	Main joint density function of random variable
$f_X(x)$	Density function of time to part failure
$h_x(\vartheta)$	Joint probability function of density sampling
$h_X(x)$	Force of mortality of time to part failure
i	Counting variable
j	Counting variable
k	Number of consecutive failed units
m	The number of items in a group, collection or list
n	The number of items in a group, collection or list
p_i	The probability of an expected event
t	Time, in appropriate units
$v(t)$	Rate of occurrence of failures at time t
x	Age of a part, measured in appropriate units
z	Z score
Cov	Variance between parameters in brackets
$CV[.]$	Coefficient of variation the random variable in brackets
D	Measurement on a rail profile
D_n	Limiting value for the K-S test
D_r	Detection factor in a RPN calculation
$E[.]$	Expected value of the random variable in brackets
E_i	Expected number of failures
$F_X(x)$	Distribution function of time to part failure
H	Heigh of a rail profile
$K - S$	Kolmogorov-Smirnov goodness of fit statistic
L	Lateral force from a train wheel
N	The number of items in a group, collection or list
\bar{O}	Mean of observed number of failures

O_i	Observed number of failures
O_r	Occurrence factor in a RPN calculation
$N(t)$	Number of failures in $(0, t]$
P_f	Probability of derailment
$R_X(x)$	Probability that part survives past age x
$R(t, t + \tau)$	Probability that a part survives past time t for an interval
\Re_m	Total number of reversals in a set of interarrival times
S_n	Step cumulative distribution function
S_r	Severity factor in a RPN calculation
T_i	Arrival time to i_{th} failure
U	Laplace test statistic
U_j	Partial derivative of a likelihood equation
U_{LR}	Lewis-Robinson test statistic
V	Vertical force exerted by a train wheel on a rail
$Var()$	Variance of parameter in brackets
$V(t)$	Expected number of failures in $(0, t]$
$Var[.]$	Variance of the random variable in brackets
X	Placeholder in the LSE method
X_i	Time to part failure
Y	Placeholder in the LSE method
Y_i	Observed value of a random variable
α	The critical value in a statistical hypothesis test
α_0	Parameter in the log linear equation
α_1	Parameter in the log linear equation
β	Parameter in the Weibull and power law equations
χ^2	Chi-Square goodness of fit parameter
η	Scale parameter of the Weibull distribution
λ	Parameter in the Weibull and power law equations
\mathcal{L}	Lagrangian multiplier
μ	Mean of a number of numerical entries
ϕ	Rotation angle
ρ	ROCOF of a homogeneous Poisson process
ρ_t	ROCOF of a nonhomogeneous Poisson process
ρ_1	$\exp[\alpha_0 + \alpha_1]$
ρ_2	$\lambda\beta t^{\beta-1}$
σ	Standard deviation of a number of numerical entries

σ_i	Standard deviation of track geometric parameters
τ	Unspecified time interval
$\hat{\theta}(x)$	Maximum likelihood estimator
θ_0	Parameter to be estimated
θ_1	Parameter to be estimated
Θ	Set of m-dimensional space
Υ	Score of a unit in a set
ζ	Military handbook test statistic

Vectors and Tensors

\mathbf{x}	Set $(x_0, x_1 \dots x_m)$
$\boldsymbol{\theta}$	Set $(\theta_0, \theta_1 \dots \theta_m)$

Subscripts

<i>geom</i>	Geometry
<i>profile</i>	Rail profile
<i>railpair</i>	Set of two rails as a single asset

Acronyms

ABR	Age based replacement
AILP	All integer linear programming
ALA	Average horizontal alignment track geometric parameter
ARE	Asymptotic relative efficiency
CBM	Condition based maintenance
CDF	Cumulative distribution function
DEA	Data envelopment analysis
FMECA	Failure modes, effects and criticality analysis
FOM	Force of mortality
FTA	Fault tree analysis
GAU	Track gauge track geometric parameter
GNSS	Global navigation satellite system
GPR	Ground penetrating radar
GTM	Gross ton mile
HAZOP	Hazard and operability study
HPP	Homogeneous Poisson process
IID	Independent and identically distributed
LSE	Least squares estimation

MLE	Maximum likelihood estimation
MPI	Maintenance performance indicator
MTBF	Estimated mean time between failure
MTT	Multiple tie tamping
NHPP	Nonhomogeneous Poisson process
PHA	Preliminary hazard analysis
PRA	Average vertical alignment track geometric parameter
PRASA	Passenger Rail Agency of South Africa
RAMS	Reliability, availability, maintenance & safety
RBD	Reliability block diagram
RMP	Risk management programme
RPN	Risk priority number
ROCOF	Rate of occurrence of failure
RPN	Risk priority number
SRPM-TQI	Adapted track quality index method
SSE	Estimated sum of squares
SUP	Super elevation track geometric parameter
TGI	Track geometric index
TSI	Track structural index
TQI	Track quality index
TWT	Twist track geometric parameter
UTS	Ultimate tensile strength

Definitions

Perway	A collection of components part of a railway track assembly, including sub-grade components.
--------	----------------------------------------------------------------------------------------------

Chapter 1

Introduction

1.1 Background

The passenger railway industry in developed countries around the world sets a standard of performance and service reliability for developing countries to follow. Railway companies in developing countries are the ‘baby children’ of the railway giants of today - French ProRail, American Amtrak and German Deutsche Bahn, to name a few. The reputation of today’s big names was not established without conquering major problems. It is important that railway companies in developing countries focus on mitigating key problems that cause the loss of company momentum rather than attempting to solve all the problems, at once. (Maluleke, 2013). Railway companies in developing countries lack the practical tools to mitigate key problems, which causes a chaotic maintenance environment where faults are not managed effectively. A practical tool needs to be developed to solve engineering maintenance problems through maintenance prioritisation. Solving high risk problems first is a sure way to improve the state of systems in an environment where not all planned maintenance will be executed.

One of the most complex maintenance environments in a typical railway company is railway perway, which struggles to recover from a sick asset fleet. Perway refers to a section of railway track. Database asset management software currently exists in many developing railway companies to manage perway, although valuable on-hand data is not well utilised. Maintenance managers need to be able to visualise and prioritise maintenance in order to get a start on a functional preventative maintenance strategy in this environment. Adding to the grass roots maintenance problem, many decision makers for perway don’t have a powerful engineering skill set to exploit trends discovered from available data. Fogel (2013) touches on the solution to these problems when he speaks about improved risk performance through effective asset management. Effective asset management is possible when decision makers have

practical tools enabling them to heal their sick fleet, with only experience at their finger tips.

1.2 Problem statement

On account of the above background, there are no tools that quantify the risk and reliability of a section of perway for the sake of maintenance prioritisation in the South African passenger railway industry. The problem is relevant because railway perway is always degrading over time due to train traffic loads. The objective of perway maintenance is to increase the reliability and availability of the perway, and improve the safety of passengers in accordance with the RAMS (reliability, availability, maintenance and safety) philosophy. Without reliable perway for trains to pass over, the railway network operations team is unable to schedule train passage effectively as there is remarkable uncertainty about the integrity of perway. Perway reliability needs to be at a high standard and this is only achieved through maintenance. Every engineering asset cannot be maintained to perfect standard due to cost and workforce limitations, which is why a strategic maintenance strategy is necessary to ensure high reliability standards for South African passenger rail.

1.3 Research objectives

The set research objectives are chronologically listed milestones leading to the primary objective, which is a reliability-based risk model. The objectives are listed as follows:

- To construct a reliability model representing the probability of successful operation of the train service from the perspective of a section of perway and populate the model using quantitative statistical failure data.
- To create a reliability-based risk model that compares the risk of severe service-stopping failures of different perway sections for the purpose of perway section maintenance prioritisation.
- To validate the developed reliability-based risk model by comparing it to an appropriate condition-based tool currently used to make maintenance decisions.

During the research process, milestones were set. These involved deadlines for surveys, interviews, data capturing, data cleaning, review of literature and actual data analysis. The crux of actual research work was reviewing of literature.

1.4 Research design and methodology

In this section, the methodology presented in Fig. 1.1 is discussed in detail. The research methodology seeks to merge avenues of research pertaining to statistics and asset management. The statistical avenue is mainly concerned with reliability analysis of perway corridors, which is boxed in Fig. 1.1 as the likelihood arm of a risk matrix. The asset management avenue is mainly concerned with the identification of perway failure modes and the calculation of average delay for each perway corridor. This avenue is dubbed severity in the figure for its contribution to a risk matrix model. A risk metric is to be computed for each perway corridor as the output of severity and likelihood calculations. The two avenues overlap because failure modes identification is necessary for the analysis of statistical reliability. The limitation of this methodology is that it applies to a specific case, namely the Passenger Railway Agency of South Africa (PRASA). This methodology was developed for PRASA to improve decision making by their maintenance engineers. As a result, data and available information was extracted from PRASA, therefore the results arising from the methodology can best be applied within this specific company.

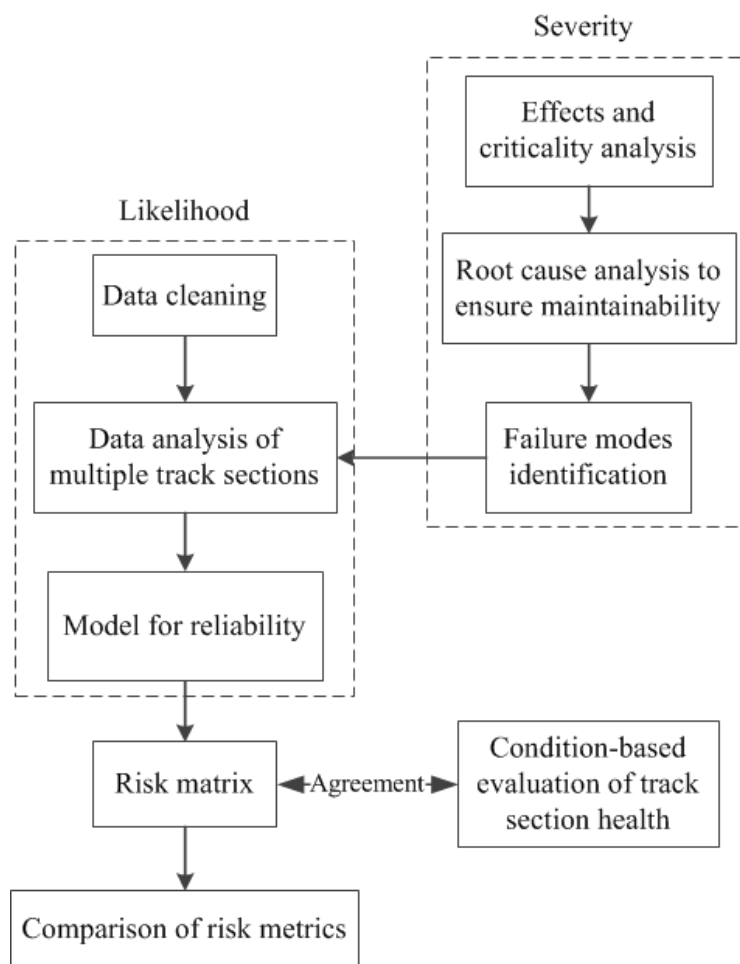


Figure 1.1: Research methodology for the development and validation of a reliability-based risk perway section maintenance prioritisation model.

Chapter 2

Literature Review

The purpose of the literature analysis is to discover the best techniques available to satisfy the primary objectives. The literature is divided into sections of which each contributed to the completion of the steps in the research methodology. The review starts with a general outlook at the given industry. It then focuses on supporting methods in order to generate the primary reliability model. The statistical basis of the reliability model is discussed afterwards and, finally, the sections are brought together to discuss risk analysis.

2.1 Infrastructure

The infrastructure functional division in a passenger railway company is interested in a high level of preventative maintenance necessary to ensure safe operation of trains. Preventative maintenance in this context is maintenance scheduled during off-peak operating hours. In comparison, corrective maintenance is the repair of sudden component failures, which often delays the train service. Infrastructure has a responsibility towards the operations department, which explains the importance of understanding the effect of these departments on each other. The operations division of a railway service company encompasses scheduling, managing capacity and measuring the performance of the systems in operation. Infrastructure and operations divisions affect each other when infrastructure faults cause delays and speed restrictions. Maintenance of infrastructure can interfere with the operations schedule of running trains, which has an opportunity cost related to under-utilised railway assets. In a similar way, capacity issues in a train network puts pressure on infrastructure to expand existing traffic routes. Performance indicators developed to monitor rail network operations performance could be better utilised for infrastructure management and is therefore of research interest.

2.1.1 Railway track maintenance strategy

The established method for large scale railway track maintenance in South Africa is condition-based maintenance, using a track geometry correcting rail car (Zaayman, 2011). This sophisticated IM2000 track car is driven along a circuit to determine poor sections of track by optical measurement and statistical analysis. Standard deviation of track geometries from the design conditions are used to construct a Track Quality Index (TQI). Poor track sections are thus separated from good track sections which results in a maintenance schedule. In addition to this, planned maintenance is conducted using inspection trolleys. Perway incidents are reported to track inspectors who send out maintenance teams to conduct corrective maintenance.

Minsili *et al.* (2012) are interested in the main causes of railway track geometry deterioration and derailment. In Cameroon, this is caused by ballast degradation. They have focused on preventative maintenance by acknowledging their number one failure cause. They developed a ballast renewal strategy which allows for improved long-term health of the track network, using FEM and a ballast inspection method which grades ballast by a deterioration index. The effectiveness of this method could pave the way to a movement in the direction of preventative maintenance for rail track and mapping of railway assets. Oyama and Miwa (2006) realised the potential for improvement of maintenance scheduling from a disorganised schedule to a schedule following an optimal route. Their degradation and restoration model makes it possible to know where and when maintenance is necessary. Equipped with this information, an all-integer linear programming model (AILP) was used to optimise the route of a multiple tie tamper (MTT), which is a perway maintenance machine. Cost was one of the decision variables in the problem. Higgins (1998) considered a train operations schedule and minimised the number of times that scheduled maintenance should clash with train operation. A tabu local search optimisation method was used to process the large amount of decision variables in the problem. This shows that track maintenance not only needs to have optimal routes for maintenance but also specific times for maintenance, in accordance with a train operating schedule. An 8% reduction in interference delay was achieved on a train schedule and 7% in maintenance completion times.

Currently used railway maintenance methods, in South Africa, require improvement and new technology. Researchers are interested in new preventative maintenance methods that eliminate wasted cost spent maintaining assets that are not critical to the success of network operation. Track geometry and ballast renewal are highlighted due to their capacity to eliminate perway failures through intelligent maintenance. New technologies for ballast and track renewal are evolving from this concern. Although preventative maintenance

seems like the answer, improved maintenance activities are not the complete answer to improved maintenance. This is because the scheduling of maintenance is equally critical to the success of a maintenance strategy as Oyama and Miwa (2006) verify. If the maintenance strategy prevents unplanned corrective maintenance then operations will incur less schedule interference, in addition to the successes experienced by maintenance departments. Railway perway maintenance contains age-based replacement, condition-based perway renewal and preventative maintenance through regular inspections and repairs.

2.1.2 Delays and speed restrictions

The press in South Africa reported that the Railway Safety Regulator issued a warning and improvement target to Metrorail Western Cape, in March 2014, about infrastructure safety compliance (Regulator, 2014). Speed restrictions of 15km/h were imposed on lines outside the Belville and the Cape Town stations as the infrastructure was declared to be in an unsafe condition. In 2014, the UK's Network Rail under-spent £1.2bn on maintenance (Topham, 2014). Infrastructure faults caused a 1-5% increase in missed punctuality targets. This increase in delays was the reason behind a £53 million fine by the Office of Rail Regulation. Sudden failures cause a disruption to regular service as trains are often delayed while faults are being fixed, which is the case for Network Rail. Another study by Börjesson and Eliasson (2011) shows that passengers care more about the length of delay and the risk of delay than the average time that trains are delayed. This is mainly due to the railway servicing business routes where time efficiency is paramount. Longer delays are mostly caused by unexpected infrastructure failures. Thus, it is difficult to make quick operational adjustments where no transfer options exist for passengers.

From the discussion, one finds that infrastructure maintenance and operations are linked through the delays and speed restrictions, which cause unreliability in the railway service. These negative effects on operations could be better controlled with perway maintenance tasks that improve the reliability of the larger system.

2.1.3 Capacity problems

Railway infrastructure is linked directly to the capacity of a rail network, which is governed by the size of the locomotive fleet, the extent of infrastructural development and train scheduling. During peak hours for a passenger train company, more trains will be in service than any other time and infrastructure failures could have an effect on train delays across the fleet. Shcherbanin (2012) investigated Russia's railway problems and highlighted capacity bottlenecks as one of the most significant problems. In a similar context, PRASA has single-carriageway train tracks between Stellenbosch and Cape Town, which causes

major delays in the event of a track failure. From these practical problems, a conclusion is drawn that high traffic single-carriageway tracks have to be maintained at a higher standard, with a higher reliability than other track sections. Gevert (2007), another researcher, encountered capacity issues that led to the expansion of Brazil's Carajas Railway. Solution methods included doubling up carriageways, at a huge capital expense and increasing train length. Often low-cost strategic solutions are overlooked and instead, solutions to increase capacity involve large infrastructure capital expenditure.

Infrastructure research should enable a higher standard of track maintenance at a lower cost, thus alleviating capacity issues by a quantifiable amount. In this way, trains will be able to travel faster and more safely. The inhibitor of such methods is the poof of their success as they encapsulate a high risk and a high reward scenario.

2.1.4 KPIs

Key performance indicators (KPIs) are often used in the railway operations division to track specific areas of performance and to compare and improve progress. This tool has been applied to infrastructure maintenance and infrastructure assets but the tool lacks traction to stimulate cultural changes in organisations such as PRASA.

Åhrén (2005) provided a definition of a key performance indicator. A 'performance indicator' is a measure capable of generating a quantified value to indicate the level of performance taking into account single or multiple aspects'. During the day-to-day management of a maintenance environment for railway infrastructure, these performance indicators can be used as a guide-line to ensure that valuable improvements are being made. Banverket (Swedish rail) used maintenance performance indicators (MPIs) that affect infrastructure, namely, train delays due to infrastructure, number of train disruptions due to infrastructure, capacity restrictions, markdowns in current standard (speed restrictions), total number of urgent inspection remarks and track quality index (Åhrén, 2005). These indicators are a guideline for research on improvement areas of infrastructure. Stenstrom *et al.* (2013) used a link and effect model to convert railway business objectives in to KPIs that were then used in a case study. This model was finally implemented at the Iron Ore line in Sweden as a standard company tool. This method allows for useful captured data to be converted into change management. The case study presented is a testimony to the success of KPI implementation in a maintenance environment, especially by engineering managers who are able to encourage buy-in from employees that is enough to initiate cultural change.

2.1.5 Safety

Infrastructure failures have a safety impact on train passengers, the most severe of these failures being train derailment. Infrastructure failures have an effect that extends beyond the consumer, causing injury to company workers as well as motor vehicle users who interact with the rail system. Financial statements of established railway companies declare safety as one of their key performance areas. Although safety can be interpreted as crime related, the same threat to life applies in an accident scenario. This shows that the reputation of a company according to both the customer and employees is a big part of company success.

Evans (2011) conducted an investigation of fatal train incidents in Europe between 1980 and 2009. Infrastructure was the second highest cause of collision for seven out of nine countries and it was the fourth highest cause for two of the countries. There were a total of 277 fatal train collisions during this period. Kyriakidis *et al.* (2012) identified infrastructure technical failures as a key precursor to railway accidents. They developed a methodology which eliminates accident precursors for the sake of prevention. This methodology was applied to eighteen major metro's that have emerged in the last decade. To mitigate precursors and reduce accident risk, it was decided that investments in infrastructure would be an important solution. Reliability improvement of railway was listed as another pro-active solution in the infrastructure domain. American rail incidents and deaths were investigated by Liu *et al.* (2011), who focused on perway failures. The number one cause of derailment was broken welds and rail sections, followed by track geometry defects. These failures modes disrupt operations as well as present safety hazards and cost complications to infrastructure maintenance. Infrastructure failures caused more derailments than rolling stock, indicating that infrastructure health is a high safety priority for railway companies. It was proven through statistical analysis that derailment risk decreases as the condition of the track improves. This is an important discovery, which indicates that there is a relationship between probability of failure and track condition. This opens up an avenue of research for railway infrastructure management that could see better established links between track condition and track reliability.

Safety in the railway environment has a strong connection to the way tracks are maintained. Indeed, a car without a road has a safety problem but the concept can now be understood in terms of railway environment. Track condition and certain modes of degradation increase safety concern and with thorough research, these problems can be better mitigated and controlled.

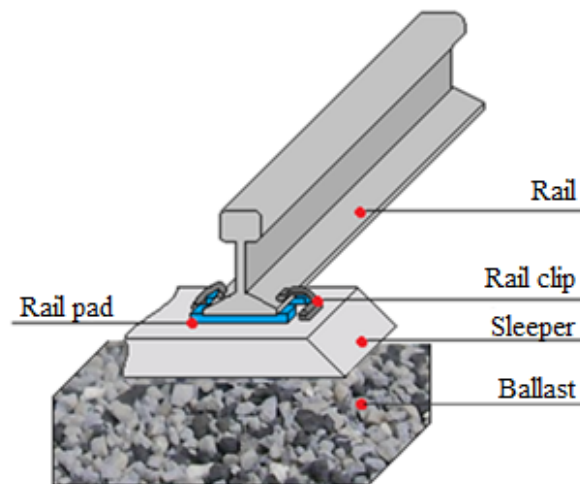


Figure 2.1: Simplified rail profile for a section of perway.

2.2 Reliability of perway components

According to (Modarres *et al.*, 2009), ‘reliability is the ability of an item (component or system) to operate under designated operating conditions for a specified period of time or a number of cycles. The ability of an item can be understood through probabilistic or deterministic methods’. In this way, reliability can be quantified at a component level or an asset level. The component level is addressed in this section. The definition of reliability can be applied to each component of the perway to ensure that it is reliability that is actually being measured. Often indexing methods or condition-based methods lack the probability of failure aspect that is an important part of the reliability definition. Deterministic approaches are better used when statistical data is not available. The definition is broken-up so that the reader can understand what is being considered for the reliability measure of each component:

- Probability component
- Time component
- Operating conditions
- Failure mode

It is pertinent to give a definition of perway as this is the primary unit of discussion for the research. A section of track is called ‘perway’ for the purposes of this investigation. Perway is defined as a collection of track components: ballast, sleeper, rail pad, rail clip and the rail itself. Fig. 2.1 presents a section of perway. All perway failure modes except block joint faults (signalling) exist

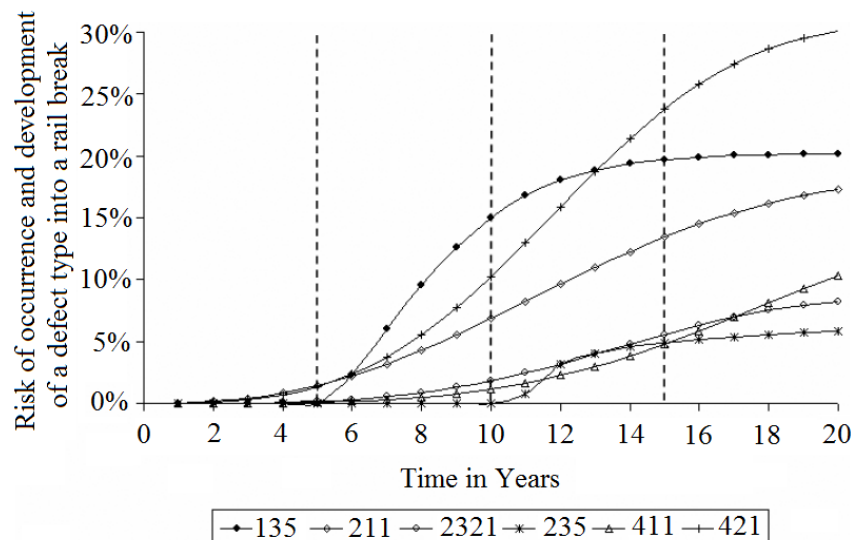


Figure 2.2: Risk of a defect type developing into a rail break for progressive age of rail (Kumar *et al.*, 2010).

due to their danger towards train derailment. Often with these failure modes, the train can still pass but the risk of derailment is so severe that the system is considered to be in a failed state. As such, probability of derailment is used as a benchmark for reliability. In literature, probability of derailment due to specific track component failures can be quantified. Keeping in mind what has been said about derailment probability, reliability for certain individual track components can be calculated statistically as well as the discussed estimation from derailment conditions.

2.2.1 Reliability of rail

Kumar *et al.* (2010) created a risk measure for rail fracture to determine the best time for different kinds of maintenance during the life cycle of a rail. Rails of the same failure code were grouped together and failures were recorded from time of installation. The risk measure is made up of a probability of occurrence component and a severity component, which was determined by experts. The probability of occurrence component is determined by assuming a repairable system and analysing the failure data, thus generating reliability measures per failure mode. Fig. 2.2 presents the risk measure for different failures codes as the rail progresses in age.

2.2.2 Reliability of sleeper

Zhao *et al.* (2007) discuss reliability analysis for railway sleepers using a failure model. They developed a reliability model for a cluster of sleepers with

consecutive failed sleepers, using k-out-of-n principles. The failure mode for a sleeper considered here is related to cracks, broken-off chunks, rail seat damage or complete fracture. A Weibull distribution is an exemplar statistical model to describe the distribution of sleeper failures. The reliability of an individual sleeper at time t is presented (Zhao *et al.* 2007, eqn. 2.1).

$$R(t) = e^{-\lambda(t/\eta)^\beta}, \quad \lambda, \beta > 0 \quad (2.1)$$

Where β and η are shape and scale parameters respectively. These parameters are determined by the maximum likelihood parameter estimation technique. The reliability, $R(T_o + \tau)$, of a single sleeper at time $t_0 + \tau$ and its probability of failure, $f(T_o + \tau)$, given that it is functional after inspection at time t_0 may be given by

$$R(T_o + \tau) = R(T_o + \tau)/R(T_o)$$

$$f(T_o + \tau) = [R(T_o) - R(T_o + \tau)]/R(T_o)$$

The above set of equations allows a sleeper reliability prediction for a future time. The further in time the model extends, the more inaccurate it becomes, which explains why discretion is advised. It is further noted that the effect of the rail, rail pad and ballast condition on sleeper reliability has been accounted for by the nature of these components interacting with a sleeper, adding to the cause of failure. The impact of these components is assumed to be random.

2.2.3 Reliability of ballast

Nurmikolu (2012) conducted condition assessment of ballast and substructure and identified GPR (Ground Penetrating Radar) as a real-time tool to monitor ballast condition. From this research, a failure mode for ballast is determined and over time, a statistical basis can be developed for ballast failure. Ballast failure is the percentage ballast material that passes through a specific sieve opening size that is more than specification, according to track maintenance personnel. Silvast *et al.* (2010) created a GPR fouling index for five meter track sections based on real-time signal data. The index data for five meter sections can be averaged over 200 meters, in-line with the reliability model. A reliability value in the time domain is generated for this data by fitting observed failures to a known distribution. Nurmikolu (2012) suggests taking samples from beneath the sleeper edges as this is where degradation is most concentrated, providing a conservative reliability estimate. In other words, more failures will occur in the model than in reality. Sadeghi and Askarinejad (2011) give equal weighting between rail, sleeper, rail clip and ballast structural condition for maintenance, thus in a reliability block diagram, the ballast

reliability measure satisfies equality. Prescott and Andrews (2013) use Petri Net modelling to capture the effect of ballast maintenance on rail geometric parameters. They stipulate that based on track geometry measures, ballast maintenance is conducted. This method is interesting because it seeks to prevent the cause of failure, rather than patch up the problem. Al-Qadi *et al.* (2008) use GPR methods to detect ballast fouling, which solidifies the notion that GPR is the future technology for ballast maintenance.

2.2.4 Reliability of rail pad

Rail pad failure cannot directly cause system failure but it can be a contributor leading to system failure. As a train drives over the track, rail clips deflect due to vertical forces and rail pads provide damping to mitigate the severity of cyclic loading on the rail clips. Rail pad failure can be characterised as the deterioration of dynamic characteristics below the design specification. This wear and tear can be attributed to fatigue which increases with number of train load cycles. Remennikov *et al.* (2006) test the dynamic characteristics (stiffness and damping) of rail pads in a laboratory environment using a direct testing method. During the tests, ton hours of loading were increased and samples were extracted from the loading chamber at different time intervals. For passenger rail, it would be more beneficial to measure installed hours rather than ton hours for a reliability model as passenger rail load varies according to passenger load, unlike freight rail. Spot tests of rail pads would need to take place at each break section to account for the effect of local conditions on the deterioration of dynamic properties. Infrastructure databases will have the installation time for each rail pad and the lifetimes could be calculated from this. Woo and Park (2014) used an accelerated loading method to evaluate the change in rail pad properties. Heat cycles were applied to the rail pads with $0 \sim 95kN$ loadings. Rail pad thicknesses decreased with increasing load frequency and pad displacement increased for a specific load case. The elasticity of the pads increased when heat was applied. Arrhenius curves were fit to dynamic load data to estimate trends for rail pad degradation with time. Although this method has validity, the real test case as presented by Remennikov *et al.* (2006) is preferred over the theoretical model (Woo and Park, 2014).

2.2.5 Reliability of rail clip

The reliability of an individual rail clip is calculated by its ability to sustain operation according to design specification. The failure mode of an individual clip would be fracture, absence of the clip from its installed position or the deterioration of the dynamic properties. The choice of the selected failure mode is further solidified by Prasad (2012), who show failure modes for the South African Pandrol e-clip under simulated load cases. For simplicity, it is better to declare a clip as either 100% reliable or totally unreliable. In this way, it

would be easier to track and process it on a system level. In contest to the chosen method, Mohammadzadeh *et al.* (2014) pose a fatigue life reliability method for the determination of individual rail clip reliability. This method applies rain-flow method and Palmgren-Miner linear damage rule for crack nucleation life and Monte Carlo simulation for a first order reliability method estimation. The approach by Mohammadzadeh is weakened as this method cannot be easily reciprocated on a real-life, large scale system.

Failures must be categorised as belonging to a straight track (tangent track) or curved track as rail clips in each of these sections experience different loading conditions. For a tangent track, train speeds are higher than curved tracks and tangent tracks do not experience horizontal force variations as significant as those on curved tracks. Rail clips on curves are expected to receive more wear-and-tear than those on tangent tracks. Marquis *et al.* (2011) state that inward cant angles like those of a curved track are more likely to fail derailment criteria than the zero cant of straight sections. This is due to the fact that the derailment coefficient for an inward cant curved section is higher. A distribution such as Weibull is used to determine the reliability of clip failures where an individual clip socket is monitored for failure.

2.2.6 Reliability block diagram

The reliability analysis on individual components is extended to account for the effect of multiple failures in combination on the whole break section. The reliability block diagram (RBD) of a section of perway comprises a logic network of components of which each has its own reliability. If a component fails (reliability zero), the system fails. The rail clip and sleeper blocks in the diagram have multiple modes of operation. These changing modes will have an effect on reliability at system level. Rail pad and ballast systems are non-critical to system failure and therefore are considered separately. It is still useful to keep track of the reliability of these components so they can be prioritised for maintenance separately. The critical system is presented in Fig. 2.3. Note that because the model consists of series connections, there is no redundancy in the model (back-ups if components fail). In railway systems, redundancy is achieved through additional structural support such as railway sleepers. If any of the sub-blocks in the RBD fails, then a train will not be able to safely pass over the tracks and the entire perway system will fail.

2.2.7 Reliability of a set of rails

The reliability of a set of rails on a railway track is quantified using the definition check list. The failure mode of the set of rails is the condition of the pair such that derailment will occur with 100% probability, according to theoretical principles. Although this does not cover every possible failure mode of track,

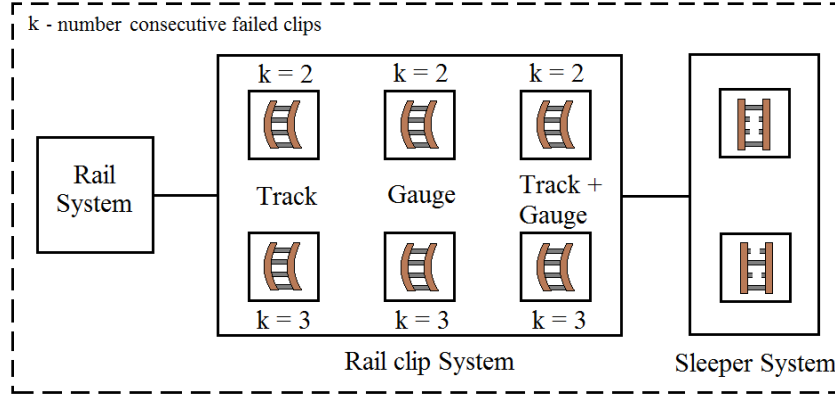


Figure 2.3: Reliability block diagram for a 200 meter section of perway.

it does cover more severe failure modes in terms of service disruption, as a first approximation of perway reliability. The time component is the design life of the rail and the operating conditions are those specified by a local railway operator. The probability of derailment, P_f due to rail geometric irregularities and rail profile wear are the components of rail reliability. These two factors are precursors to rail cracking and fracture. The self-explanatory equation (2.2) quantifies the reliability of the rail pair based on series reliability system theory. This equation is an initial estimate of the rail set system reliability. Note that a break length of 200 meters is considered for reliability calculations.

$$R_{railset} = (1 - P_{f_{geom}})(1 - P_{f_{profile}}) \quad (2.2)$$

Mohammadzadeh *et al.* (2011) used five geometric input parameters with random values to determine derailment probability from track geometric irregularities using a Monte Carlo simulation method. The Nadal criterion for derailment was considered in the formulation of the probability of failure. Nadal developed a derailment coefficient limit, L/V , which is a limit on the lateral over the horizontal forces experienced by a train wheel in operation. The equation used to calculate derailment probability is described (Mohammadzadeh *et al.* 2011, eqn. 2.3). Further detail on the method is presented in the literature.

$$P_f = E \left(I(\vartheta) \frac{f_x(\vartheta)}{h_v(\vartheta)} \right) \approx \bar{P}_f = \frac{1}{N} \sum_{j=i}^N \left(I(\vartheta_i) \frac{f_x(\vartheta_i)}{h_v(\vartheta_i)} \right) \quad (2.3)$$

Where: $h_v(\vartheta_i)$ is the joint probability function of density sampling. $f_x(\vartheta_i)$ is the main joint density function of random variable. N is the total number of tests for Monte Carlo analysis. $I(\vartheta_i)$ is the indicator function with a value of one if x is located in the failure region and zero if x is located in the safe region. Further detail on the first term of the above equation is described in the literature. The described methodology can thus be followed to determine

probability of failure based on track irregularity parameters. An input measurement from a recording rail car is necessary to use this method. The final step is the application of probability equation (2.3). The effect of non-critical perway components (rail pad failure and ballast distribution) on the probability of failure due to geometric input parameters is random and thus will not skew the calculation of $R_{railset}$.

Mohammadzadeh and Ghahremani (2012) use another approach to incorporate the effect of rail profile wear on probability of derailment. As a rail profile deteriorates, a train wheel glides up the rail to an unstable position such that an impact on the wheel can cause the train to jump from the tracks. A track recording car is used to measure the rail profile parameters using a non-contact, optical method. Van der Merwe and Venter (2001) confirm the performance capabilities of the IM2000 rail car used by PRASA. The methodology for the calculation of probability of derailment due to rail profile wear is presented in Fig. 2.4. Further detail on the method and parameter definitions are presented in the literature.

2.2.8 Reliability of the sleeper system

Immediate sleeper system failure is when two consecutive sleepers fail, causing a service disruption. A more common occurring failure is when a threshold number of individual sleepers fail in a dispersed manner in a small section of perway. Reliability can be calculated from statistical data for two or more consecutive failures as well as reliability depending on how many dispersed failures currently exist in the section. The cluster reliability for four cases of k-out-of-n sleeper failures is calculated. The cluster reliability includes $k = 2$ and $k = 3$ consecutive sleeper failures that exist between sleeper cluster sections. Fig. 2.5 presents the series sequence used to calculate sleeper reliability and a recursive sequence when the system becomes more complex ($k = 3$).

2.2.9 Reliability of the ballast system

The ballast system is a non-critical component in terms of derailment and thus it is not presented in Fig. 2.3. The ballast reliability is determined by the average reliability of a group of samples in a break section. Thus, ballast reliability at the system level is simply a collection of statistical failure data from ballast failures at key points along a track section.

2.2.10 Reliability of the rail pad system

The rail pad system is a non-critical system for reliability. To determine rail pad reliability, the time the rail pads have been installed must be recorded and compared to dynamic degradation curves generated by Remennikov *et al.*

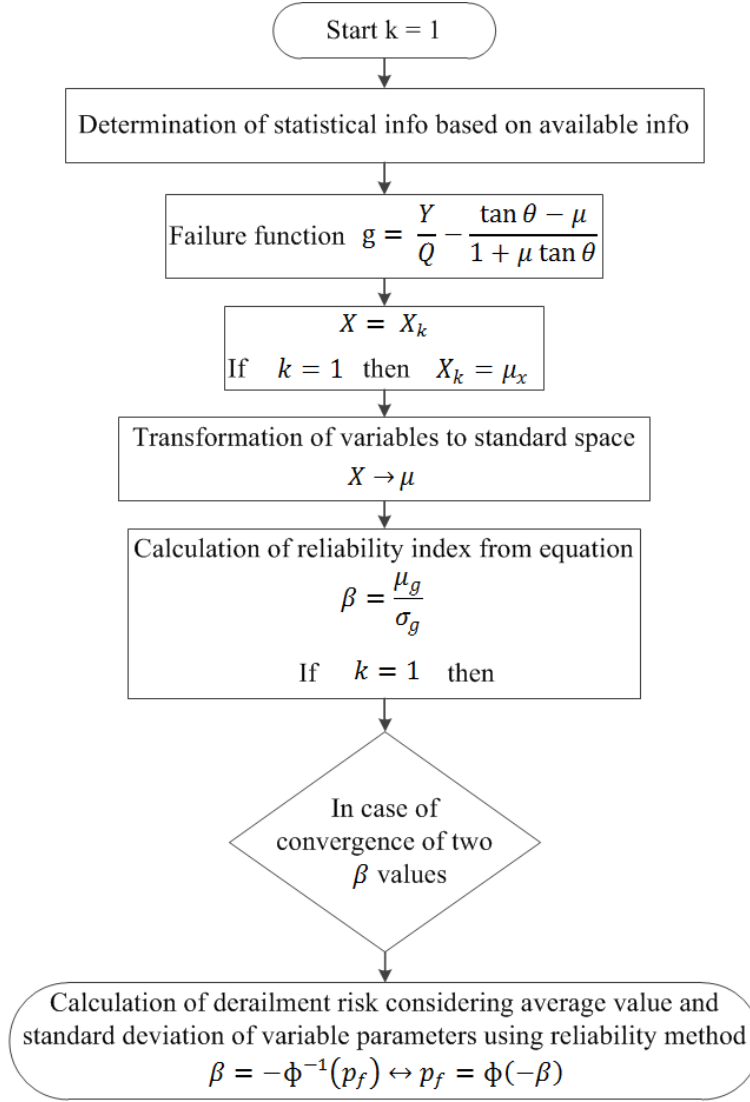


Figure 2.4: Methodology to determine train derailment probability (Mohammadzadeh and Ghahremani, 2012).

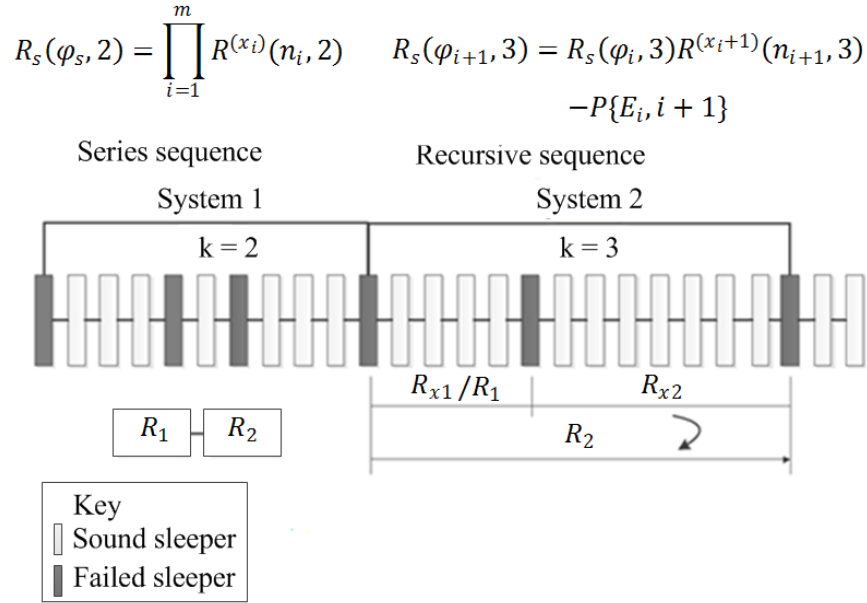


Figure 2.5: k-out-of-n reliability analysis for $k = 2$ and $k = 3$ systems (Zhao *et al.*, 2007).

(2006). Spot tests should be conducted on rail pads that are removed from each break section at specified times after installation to check the effect of local conditions on the deterioration of the rail pads. Failure records will then be adjusted according to condition data. This is a reliability estimation method based on operating conditions. A failure database will be built up using this method and from this, reliability can be determined.

2.2.11 Reliability of the rail clip system

A train can derail when a number of consecutive rail clips fail. The reliability of the system will depend on the specific location of rail clip failure and how many consecutive clips have failed. Similar to the sleeper system, reliabilities are averaged over the break section and act as part of the series critical system of perway. The derailment mechanism that rail clip failure can cause is called ‘rolling track failure’ (Iwnicki, 2006). This occurs when the vertical and lateral derailment forces from the rolling stock overcome torsional stiffness and rail clip restraining forces causing the track to pivot on its track side edge. The failure mode for this phenomenon is presented in Fig. 2.6. The rail rollover failure mode is further confirmed by Greve *et al.* (2014). They measure pressure experienced by the rail seat for increasing coefficients of derailment. The findings reveal high pressure on the track side of the rail seat with little pressure on the gauge side for high coefficients of derailment. This indicates an outward rolling moment for the rail. For rail rollover failure modes, the rail experiences poor resistance to the moment forcing the rail over due to wear

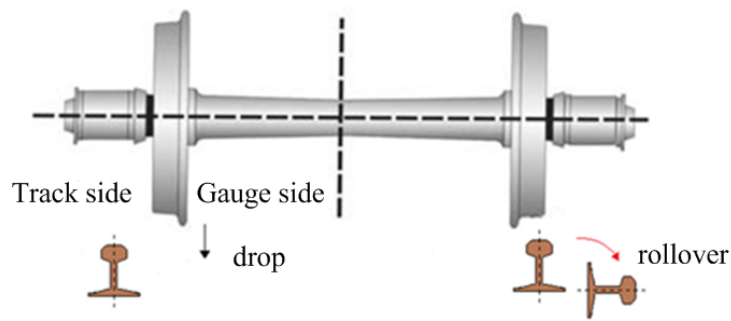


Figure 2.6: Rail rollover failure mode (Iwnicki, 2006).

and tear, states Marquis *et al.* (2011). To illustrate the causes of rail rollover, The researcher developed a simplified beam model in Fig. 2.7. This figure provides support to solidify the argument that consecutive clip failures have a drastic effect on probability of derailment when compared to dispersed clip failures.

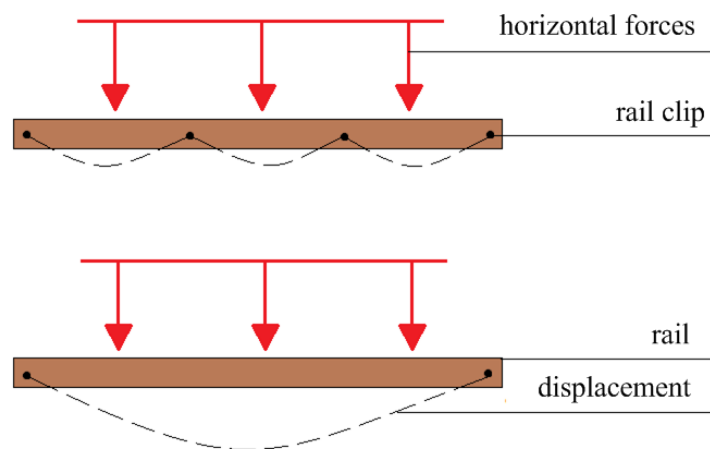


Figure 2.7: Top view of train loaded rail with rail clip supports.

Fig. 2.7 illustrates that with consecutive failed supports, the rail is more likely to bend. This bending causes horizontal forces to act at the top end of the rail, initialising a tipping motion. This argument is supported by the US Dept. of Transportation (2011) which shows that for three consecutive clip failures, the gauge widening resistance limit of a track section is exceeded. This is considered to be a system failure. Research on k-out-of-n sleeper failures by Zhao *et al.* (2007) can be used to determine the reliability of a rail clip system.

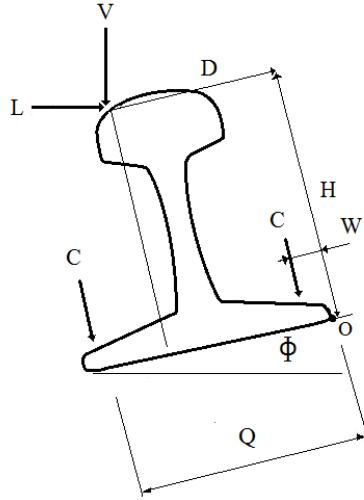


Figure 2.8: Moment diagram of rail section with train forces and clip forces (Iwnicki, 2006).

This research is consistent with observed reliability theory for consecutive k-out-of-n: F systems (Elsayed, 2012). Six possibilities of failure are considered, which are presented in Fig. 2.3. Train derailment is most likely for consecutive failed track and gauge clips, less likely for gauge side consecutive failed clips and least likely for track side consecutive failed clips. The difference between each of these failure modes in terms of applied forces is understood by the moment equation around the rail edge for rolling rail derailment, equation (2.4). Fig. 2.8 presents a moment diagram of a rail section under loading from a train, which relates to equation (2.4). The C variable from the figure represents reaction forces from rail clips when loaded by a passing train.

$$M_o = L'H - V'D - CW - CQ \quad (2.4)$$

$$L' = L\cos\Phi - V\sin\Phi$$

$$V' = V\cos\Phi + L\sin\Phi$$

The researcher proposes that a sensitivity analysis for different failure modes be conducted. In other words, the effect of each of the clip forces on the moment in equation 2.4 must be quantified by changing one variable while fixing the rest of the system. The critical moment for rollover is exceeded when L/V is large and C is zero. From the sensitivity analysis, an estimate for a coefficient can be determined that will be applied to the reliability for consecutive track and gauge side failures respectively. The reliability will be lowest for consecutive track side and gauge side failures, higher for track side failures and highest for consecutive gauge side failures.

2.3 Track quality index

Large engineering systems like the railway networks, land to sea material conveyor networks and power stations all have something in common. They have large engineering assets that need to be managed at a system level. Management is necessary because these assets are capital intensive and cannot simply be discarded and replaced when they fail. Scoping in on the perway assets of a train company, it is perceived that maintenance accounts for a large portion of the total expenditure in the life of the system (Esveld, 2001). Due to the importance of the infrastructure division of a passenger railway network and the contribution of maintenance as an operating expense, research into improved maintenance methods and practical maintenance tools is being undertaken all the time.

Improved maintenance methods are being researched by PRASA and Transnet, the railway giants of South Africa. PRASA is interested in practical tools that can improve the reliability of track sections through maintenance. A current condition based tool of particular interest is track quality index (TQI). TQI encourages maintenance to be conducted by priority, which is necessary to eliminate current maintenance backlogs. TQI grades track sections according to their track quality, which is calculated as the standard deviation of track geometry measurements from the design conditions. The researcher has investigated separate components of railway perway to accurately quantify the reliability of a section of perway. Now, the researcher understands the condition of the perway as a whole through TQI.

2.3.1 Track quality index definition

Track geometry changes over time due to subgrade and ballast shift, failure of rail clips on the track and rail/wheel contact exerted by the train. This change in geometry is called track irregularity. Track irregularity is defined by five geometry parameters, namely, twist (TWT), track gauge (GAU), super elevation or cross level (SUP), average vertical alignment (PRA) and average horizontal alignment (ALA) (Zaayman, 2013). Mean value measurement of track irregularity is conducted on 200 meter break lengths of track by the Plasser IM2000 recording car, instituted by Transnet. This car uses optical measuring techniques to measure geometry deviation data and is presented in Fig. 2.9. The track irregularity of a break length is quantified by Track Quality Index (TQI), which is the sum of standard deviations of the five track irregularity parameters. The South African method for TQI, as discussed, is simple when compared to that in use internationally such as the TGI index used in India (Sadeghi and Askarinejad, 2010), which has weighted components. China's track irregularity system has more considerations than that of South Africa, with 7 geometric parameters considered (Xu *et al.*, 2011). An



Figure 2.9: Plasser IM2000 geometry measuring rail car (Zaayman, 2013).

equation was developed for TQI (Xu *et al.* 2011, eqn. 2.5) which is consistent with measuring techniques used by PRASA and Transnet,

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (c_{ij}^2 - \bar{c}_{ij}^2)}$$

$$TQI = \sum_{i=1}^5 (\sigma_i) \quad (2.5)$$

$$\bar{c}_i = \frac{1}{n} \sum_{j=1}^n c_{ij}$$

where σ_i is the standard deviation of measurements of the i_{th} track irregularity parameter at markings in the break length, c_i is the average of measurements of the j_{th} track irregularity parameter, c_{ij} is the measurement of the i_{th} track irregularity parameter at the j_{th} marking in the break length and n is the number of measured markings in the break length. The higher the TQI is, the worse the condition of the track becomes. A threshold value of 7.5 exists for high speed passenger lines at PRASA for the purposes of safety.

2.3.2 Using TQI to make maintenance decisions

Classically TQI is used as a prioritisation tool for track lifting, levelling and tamping. Ideally, railway perway can be explained as an elastic system that deforms according to large train traffic loads and self-corrects. In reality, fatigue causes the deformation of perway until such a point as speed restrictions

Table 2.1: Key contributors leading to track geometric defects (Sadeghi and Askarinejad, 2009).

Geometry parameter	Effect of track structural conditions on geometry parameters			
	Rail condition (%)	Sleeper condition (%)	Fastening condition (%)	Ballast condition (%)
Gauge	24	21	45	10
Profile	10	24	18	48
Alignment	16	33	14	37
Twist	6	36	3	55

are imposed on a line for safety control of a deteriorating perway section. Large machines have been in operation for years which correct the geometry of a track as close as possible to design conditions. The exercise of lifting, levelling and tamping is the deconstruction of a perway section, the lifting of components and the repacking of these components into ballast.

Correcting track geometry through tamping does not necessarily solve track geometry problems but can sometimes only provide a temporary fix. Sadeghi and Askarinejad (2009) group track structural components as either belonging to rail, sleepers, fastenings or ballast in an investigation of the Iranian railway network. According to Sadeghi and Askarinejad (2009), the condition of each of these four components has varying effects on the standard deviation of track geometric parameters. Determining the effect of each component on standard deviation of track geometry will allow better informed maintenance decisions from data collected by a track recording car. Once the components were grouped, the defect density of each component type was calculated by dividing the number of defected two meter track sections by the total number of sections in a given sample. An overall defect density is calculated, with a filter coefficient for severity of defects as either low, moderate or high. Trends between defect density of components and standard deviation of track geometry were plotted for measurement on the same track section. The stronger the trend, the larger the contribution of the particular component to the geometry parameter. The trend strength indicator was converted to a percentage scale and the results of the analysis are presented in Table 2.1. The analysis indicates that track fastener condition is the primary contributor to track gauge deviation. Rail profile wear is caused by ballast deterioration. Track alignment deviates due to sleeper and ballast degradation and track twist, although weakly correlating to component failure is caused by ballast and sleeper deteri-

oration. The results of the discussed analysis are evidence towards the creation of a maintenance decision diagram based on track geometric irregularity.

2.3.3 TQI prediction vs reliability prediction

As with reliability, generating TQI's future predictions for the sake of maintenance planning is possible. Sadeghi and Askarinejad (2011) developed a track degradation model which determines how track condition indices change with time. The TGI (Track Geometric Index) similar to TQI was evaluated against time, train speed, TSI (Track Structural Index), loading and initial TGI. During tests on the Iranian rail line, all variables were fixed and only one of the discussed variables was allowed to change. The model ultimately produced a graph which presents the relationship between TGI and time. TGI and TQI have the same inputs but the calculation of each parameter is slightly different. The same methodology to produce this model can be applied to a South African rail context to produce a model for changes in TQI over time. An average estimate of TQI for different categories of tested track sections is presented in Fig. 2.10. Note that the decreasing TGI trend (negative) is opposite from expected when considering TQI. This is because with TGI, the geometry irregularity parameters are inputted into a function with a negative exponential before summing together for the TGI. Detailed definitions can be reviewed in the literature.

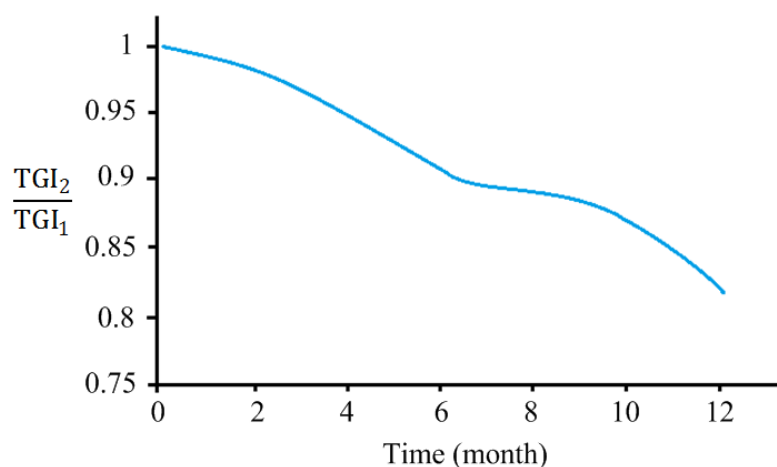


Figure 2.10: Averaged estimate for change in TQI between maintenance intervals (Sadeghi and Askarinejad, 2011).

Xu *et al.* (2011) investigated a short-range prediction model for track quality index. A track irregularity prediction technique, SRPM-TQI was used in accordance with historical TQI calculated from waveform data. This approach

applies the least squares estimation method (LSE) to develop a prediction model for each break length of track. LSE produced errors of less than 8% for sixty break lengths of track. Abnormal deviations occurred when tamping or levelling operations were in progress during measurement intervals. Two methods were discussed to improve least squares predictions when these operations were in play. Sadeghi and Askarinejad (2010) and Xu *et al.* (2011) juxtapose two methods for future TQI prediction. Sadeghi and Askarinejad (2010), on the one hand, used a scientific method of physical testing where all variables, except one, are fixed in such a way that the relationship for changing a single variable could be determined. Xu *et al.* (2011), on the other hand, used a statistical LSE to predict future TQI readings. The method by Sadeghi and Askarinejad (2010) is more advantageous to calculate future TQI predictions than that used by Xu *et al.* (2011). This is explained by the fact that the TQI prediction method used by Sadeghi and Askarinejad (2010) involves controlled tests that eliminate outside interference. Although the statistical analysis by Xu *et al.* (2011) involves more rigorous techniques, Sadeghi and Askarinejad (2010) provide a best fit for data from which predictions were made and the analysis by Sadeghi and Askarinejad (2010) is more comprehensible.

It is suspected that TQI has a relationship with perway reliability as TQI represents different states of the track. It has been uncovered that these states have threshold values which indicate failure modes for certain values of TQI. Like reliability, TQI provides a means to maintain the track. Fig. 2.11 presents how the TQI is broken down into components which are inputs to the probability of derailment equations (due to geometric irregularities). Fig. 2.11 proves that TQI is a function of σ_i and P_f is a function of σ_i therefore a relationship must exist between TQI and P_f .

2.3.4 Further considerations for TQI and perway reliability

It was uncovered that TQI and perway reliability are related as they both change according to the same input parameters. As TQI is a linear equation of standard deviations, it is easy to monitor the effect of changing input parameters on the TQI output. The change of input parameters on reliability is not easily determinable due to the dynamic nature of the perway environment. One can expect that for extremely poor TQI, the reliability of a perway section will reflect a comparable level of degradation. A reliability block diagram was developed which enables the calculation of perway reliability. Prioritising maintenance based on failure statistics is a sure way for maintenance to reduce the number of track failures. Condition-based maintenance from tools such as TQI can be coupled with failure statistics to harness the power of reliability centred maintenance. This combination of methods is imperative for strategic

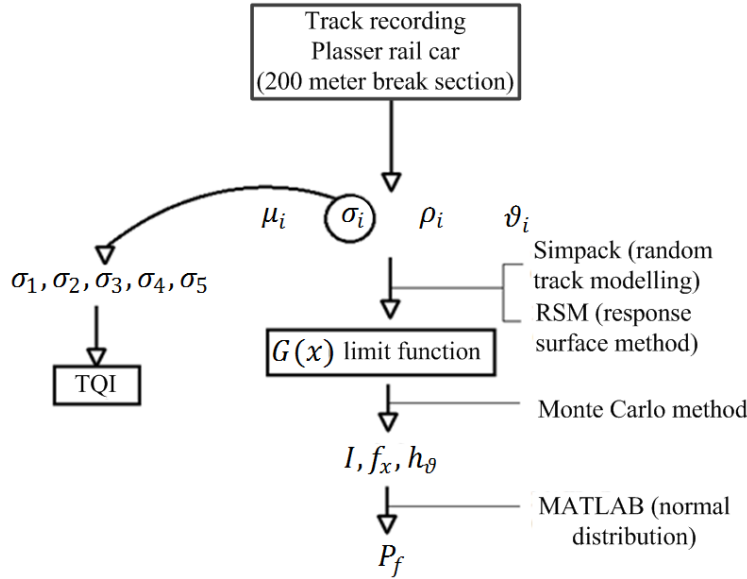


Figure 2.11: Relationship between TQI and rail derailment probability due to track irregularity.

cost management of maintenance and the improvement of safety in the railway industry.

2.4 Effects and criticality

In order to understand perway failure modes and their effect on the passenger railway service, effects and criticality methods were investigated for their applicability. Failure modes, effects and criticality analysis (FMECA); pareto analysis; cause-consequence analysis; hazard and operability study (HAZOP) and preliminary hazard analysis (PHA) were compared for applicability in determining perway faults and their criticality.

PHA can be compared to HAZOP in that they both place priority on hazards rather than service failure. The most severe kind of hazard is one that causes death of personnel. The weakness of PHA is that it is accident focused. The methodology of PHA reveals harm to personnel, without placing priority on the expected failures of the system. HAZOP identifies deviation from design operating conditions using ‘key words’, which stimulates creativity in the team analysing operating hazards. The methodology is not favoured as it requires a team of experts to identify hazards and the solutions to hazards do not necessarily solve system failures. A cause-consequence analysis evaluates each consequence of an undesirable event and evaluates probabilities of each of these occurrences. This method is thorough but requires the analysis of

every possible consequence of a fault. The analysis provides insight into the effect of a fault on other subdivisions in railway but the detail of this analysis draws away from the primary objective of the investigation. A pareto analysis captures the maintenance problem of railway perway divisions that are unable to prioritise high risk events as first priority. The downfall of pareto is that it reduces the failure risk across the perway network and does not necessarily allow prioritisation of perway section A to perway section B for maintenance. This is important as maintenance is sectionalised due to the large land area over which the assets lie. FMECA is the only tool which seems to allow the identification of failure modes that have a criticality that can be compared on a sectional basis. FMECA is understood in the context of industry to identify its limitations.

Failure modes, effects and criticality analysis (FMECA) is ‘an engineering technique to define, identify and eliminate system errors from the system design, process or service before they reach the customer’ (Chin *et al.*, 2009). This technique is adapted as a diagnostic tool to identify railway perway failures in current operation. The mechanism that conventional FMECA uses to eliminate system errors is the Risk Priority Number (RPN) method. The RPN measure is attractive as the researcher has identified its compatibility when coupled with reliability, as it reveals a risk ranking of failure modes that a reliability model is unable to capture. The RPN measure is calculated as (Chin *et al.* 2009, eqn. 2.6):

$$RPN = O_r \times S_r \times D_r \quad (2.6)$$

where O_r is a factor grading the occurrence of failure, S_r is a factor for severity of failure and D_r is a factor representing detection of failure. These factors each have a 1 - 10 scale as stipulated by literature. The RPN method has been scrutinised in literature for its inability to identify hidden risks. When comparing two RPN’s of the same number, two drastically different situations can be at work (Wang *et al.*, 2009). Chin *et al.* (2009) used data envelopment analysis (DEA) to determine risk priority as they do not believe that the RPN method is accurate enough for an industry wide application. Another fix for the shortfalls of RPN was realised with the weighted risk priority number evaluation (Xiao *et al.*, 2011). This method employed a minimum cut set-based algorithm to generate a new means for FMECA. Further improvements to the FMECA were realised through fuzzy logic with expert opinion (Xu *et al.*, 2002). This approach was taken because interdependencies between failure modes are often difficult to incorporate into FMEA. A similar approach was taken by Wang *et al.* (2009) who decided that too many fuzzy decisions were to be made in FMECA and reducing the number of decisions induced error. An alternative approach, fuzzy weighted geometric mean, was incorporated into the FMECA. It is argued that although many improvements have been made

to FMECA, the application of FMECA determines whether the standard or an adapted FMECA should be used. A methodology for basic FMECA in railway perway was developed from understandings gained in literature; this is presented in Fig. 2.12.

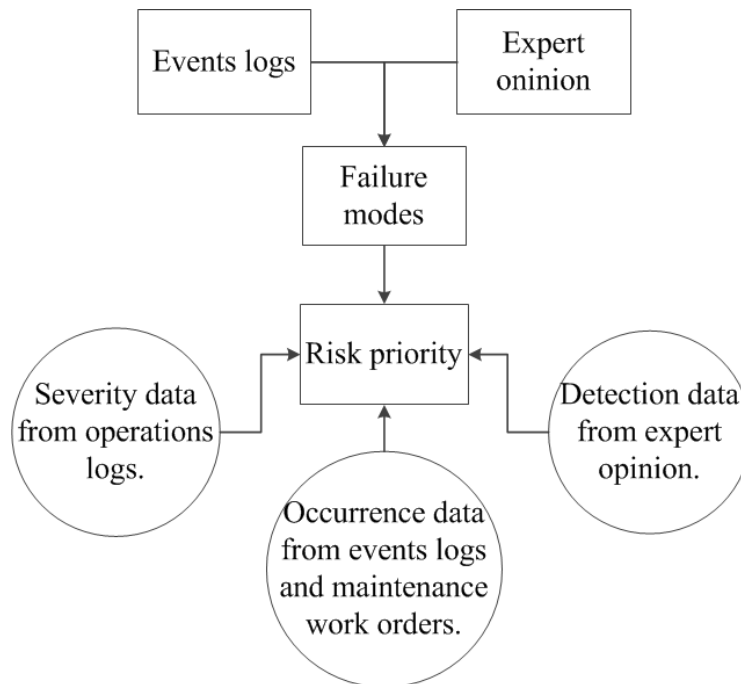


Figure 2.12: Methodology for the construction of FMECA for railway perway.

Irrespective of any scrutiny, FMECA is widely used around the world and was proven successful in a test of 100 FMECA applications in Japanese industry (Xu *et al.*, 2002). Standard FMECA has been applied to railway infrastructure in Sweden (Morant *et al.*, 2014). Here, FMECA was used as a decision support tool for maintenance of signalling systems, which is close to home for perway. The FMECA in this report was built from the standard, discussed by Wang *et al.* (2009); considering the adaptation by Sameni (2012). As FMECA has not been widely utilised in PRASA, the standard FMECA method is used as a first step towards quantifying the reliability of a section of perway. This being said, the literature provides insight into potential complexities when comparing risk for failure modes. The standard FMECA is not accurate enough to be the sole indicator of maintenance priority and the detection measures in RPN may be over-weighted in its application (Sameni, 2012).

2.5 Root cause analysis

When failure modes are identified for perway failure, it is beneficial to ask: ‘What caused the failures?’ This helps in determining which maintenance task can prevent the failures from occurring, which reduces the necessary amount of corrective maintenance. Root cause analysis methods were investigated to determine which methods would be relevant: Fault tree analysis (FTA), Ishikawa (fishbone diagram), the Five Whys, causal factor charting or root cause mapping. The Ishikawa diagram is a method designed to link cause and effects together. It is useful when a dominant negative effect exists which needs to be reduced. The limitation of this is if the effect itself cannot be easily traced to the root of a problem. For example, perway maintenance is concerned with service disruption thus if train delay is the effect then it would be impossible to identify all relevant faults through brainstorming. It is better to first identify failure modes from which problems or root causes can be identified. The Five Whys method adapted from six sigma is applicable to simple cases such as human error. It is not extensive enough to cover causes of faults as experienced in railway perway. Casual factor charting, root cause mapping and FTA all have a common thread in that they consider a sequence of events that lead to a failure. Casual factor charting is separated from the others in that it doesn’t necessarily lead to root causes. It simply identifies conditions that were critical to the occurrence of a failure. The limitation of this method is that it does not apply well to failures with multiple possible combinatorial causes as encountered in railway perway. Root cause mapping is similar to FTA except that it doesn’t rely on logic but rather on collected evidence from each potential cause of the failure. This method is more practical where field work is being conducted. A researcher can use FTA from a theoretical standing to investigate possible faults of pre-identified perway failures. The logic component of FTA helps determine priority of root causes where redundancy exists. Geum *et al.* (2009) define Fault Tree Analysis (FTA) as: ‘a means to translate a physical system into a structured logic diagram, with the aim of identifying the faults in the system and their influence on system function’. The purpose of conventional FTA is to expand on a fault and trace the logic back to the root cause of the fault. A methodology was developed for the application of FTA to the railway context, presented in Fig. 2.13. The applicability of FTA is explained.

Shalev and Tiran (2007) state that FTA is used primarily as a design tool and propose a method of condition based data to create real-time failure rates for each potential fault in the tree. This is an improved approach to the conventional failure rate estimation, which is used during the design process of a system. Their approach reveals that fault tree analysis can be used not only during the design phase but when re-evaluating systems as well. FTA is adaptable beyond the standard failures approach, to be used in other indus-

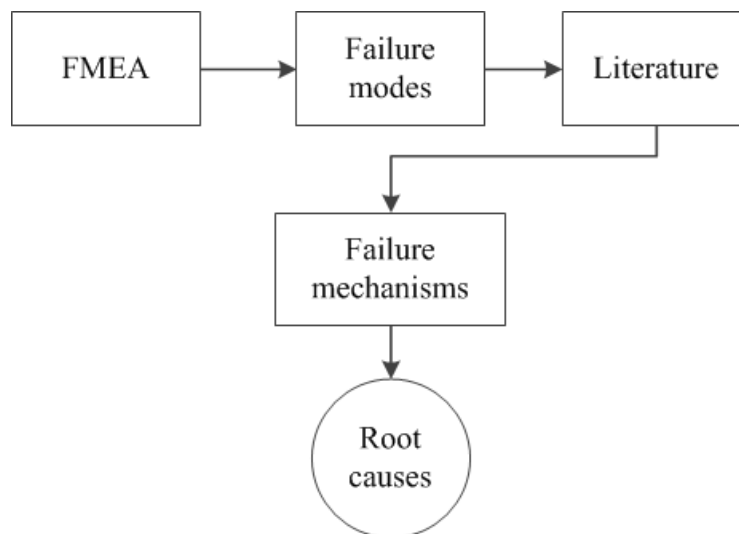


Figure 2.13: Methodology for the construction of fault tree analysis for railway perway.

tries. This is proven by alternative approaches in literature. Fuzzy fault tree analysis was created by Jafarian and Rezvani (2012), which is a useful adaptation to the standard process when quantitative analysis is not readily possible. This approach was applied to the passenger railway industry to identify root causes of passenger train derailment. The fuzzy method uses importance measures and error rates to calculate top event probability. In another adaptation, fault tree analysis is used in combination with a Markov model, to assess time independent and dependent factors together. The field of application is medical, where failure probabilities are dependent on the time the patient has been under the knife. Basic events are categorised in the fault tree and continuous time failure probabilities are calculated using Markov chains. The tree is modularised, which allows both standard tree calculation and Markov representation in the same model (Zixian *et al.*, 2011). Similar to a time continuous instance, phased mission modelling is necessary to capture certain failure instances. An example of this is the different phases in a space shuttle mission, where certain failures are critical only in specific mission phases. With this method, fault trees are modularised for the different mission phases and are then converted into binary diagrams for mathematical computation. The failure probability of each mission phase can be calculated from these diagrams (La Band and Andrews, 2004). It is apparent from the literature that, due to the logical nature of fault tree analysis, it is adaptable to different applications and fault tree concepts can be translated to equivalent concepts. Fault tree analysis is proven to be applicable to a continuous time situation and is not confined to a stagnant state. Adaptability is applied by the researcher when analysing railway perway to determine root causes of critical failure modes. Identifying root causes is simply for prevention purposes although mainte-

nance and failure rates are not needed for each potential cause in the logic tree. As complex fault tree methods were developed to determine failure rates with little information, the standard fault tree will instead be applied in this context, where failure rates are not necessary.

2.6 Track service life

According to the definition of reliability, section 2.2, life-time or service life of a system is critical to its reliability. Exceeding a component or system life-time as specified by the OEM or relevant standard could cause unreliability in that system or component as its function degrades to unusable conditions. This kind of component or system is considered unreliable. It is important to understand how service life of perway components is measured so that service life at a system level can be better understood.

Perway is primarily concerned with rails, sleepers, rail clips, rail pads and ballast therefore components can be individually decommissioned by understanding the typical service life of components and tracking their current life. In comparison to a rail, a concrete sleeper has a much higher service life (approximately 50 years) and the ballast is always shifting thus the rail is the component of primary interest for service life considerations. Rail pads and rail clips are subcomponents to rail and thus will not be able to provide clarity to help classify perway lifetime. Fatigue is the root cause of much of the component degradation which is a time and load based phenomenon. Service life of track components is generally measured by GTM (gross ton mile) as there is no reliable means to determine fatigue damage for these components (Mundrey, 2010). Once main-line rails have reached their service life limit, they are used on sidings and in yards. This is attributed to the fact that rails are often removed before the end of their useful life for safety precaution. This switch-over happens after approximately 20 years or according to a rated GTM limit. The limits are presented in Table 2.2 from the Belgian standard. The life of rail clips and rail pads vary significantly between competitors but it is important that they maintain their dynamic properties as long as possible. A service life of 20 years has been cited from Washington Metropolitan Railway Service.

With large or complex systems, deciding the age of the system is difficult as components in the system can have varying ages. Some components, such as railway ballast, shift/change and thus are a mix of components of different ages. This anomaly identifies perway as a complex system for which some sort of life-time model should ideally exist. Without thorough data for the system under consideration, it is impossible to disqualify components in the system due to their age. If immediate exclusion is not possible, degrading systems are

Table 2.2: Rail service life (Mundrey, 2010).

Rail section	Total GTM carried
60 kg/m	500
52 kg/m	300 - 350
90 lb/yd	250 - 350
75 lb/yd	150 - 350
60 lb/yd	125 - 350

Note: The service life of the rails indicated above is for standard quality rails with a UTS of 72 kg/sq mm. For rails with a UTS of 90 kg/sq mm, the service life is taken as $1_{1/2}$ times that of standard quality rails.

identified through reliability trend analysis. Typically, there will be a number of track components in a track section that have passed their specified service life. The system as a whole cannot be disqualified due to these old components as they present only a small fraction of the total system. Suppose this fraction is large and not small, it is known that rails are still used after their service life for other applications. These systems are old but not necessarily unreliable in a practical context. It is thus recommended by the researcher not to evaluate assets according to their life-time but rather assess the condition of assets to estimate useful life left. If an individual track asset is causing downtime on a track corridor then the asset should be replaced without declaring the corridor fundamentally unreliable. Old components are thus a limitation to reliability analysis and old systems should not be compared with new ones in terms of reliability. For example, a newly laid section of track should not be compared with an old section in terms of reliability but old sections can be compared to each other.

2.7 Data cleaning

Raw data is often obtained with missing data entries, repetition, omission of information and bunched data in a single entry. Efforts are being made in the academic world to provide citation for data sets to create useful data that can be re-used. The management of raw data to determine statistical trends is very important if the accuracy of the discovered trends is to be conserved. Data cleaning procedure controls the accuracy of data sorting as well as the careful omission of data entries from a dataset. Literature was reviewed to understand which methods for data cleaning are acceptable.

Riera-Ledesma and Salazar-González (2007) produced an optimisation method that highlights the minimum number of data entries that need to be changed to fulfil a new consistency criteria. This allows a previously inconsistent dataset to be validated. They use a descending search algorithm, accelerated by Benders' cuts obtained by using Farkas' lemma on infeasible sets. This produces a near optimum solution which locates the minimum number of entries to be edited and uses an input operation to restore the feasibility of the dataset. This research suggests that a level of consistency is expected in data records. Sun *et al.* (2013) analysed yield data for crops located on a hillside to clean up unreasonable outliers, distribution outliers and geographically misplaced data. The software removes data greater than 2.5 standard deviations from the mean yield as well as removing crop entries that are inconsistent with the mass of the yield per area. These operations indicate the necessity to remove outlier and inconsistent data from the set. A Global Navigation Satellite System (GNSS) was used to record geographic location of crops for planting. Passes were tracked when planting and a location error was determinable when comparing data co-ordinates to designated planting paths. This method can be applied to railway infrastructure to record maintenance events geographically when teams are in a rush and unable to record entries with the correct level of detail. Location data is helpful to provide certainty that a data entry is indeed valid. Bertossi *et al.* (2013) attempted to clean data by combining multiple entries conveying similar information into a single data entry. They used a matching dependency to do this, which is a function that combines the two data entries correctly. An operation selects given data entries with identified similarity using a search function and breaks up the entry into definite data strings. The strings are compared and duplicate data is copied to a new cell. This concept can be applied to infrastructure data to remove duplicate data as well as combine similar data entries into single maintenance actions. These data cleaning ideas converge in a paper by Van den Broeck *et al.* (2005) which focuses on data cleaning techniques. They speak about errors when conducting research, errors when entering into a database and errors when analysing the database and extracting data. They also discuss error possibilities for database entries and diagnostic tools to improve the data, which will be the stage at which data is encountered during this research. They further provide a framework for thorough data analysis as presented in Fig. 2.14. From this framework, in combination with the other methods discussed, perway maintenance data can be cleaned to an acceptable standard before statistical analysis is conducted.

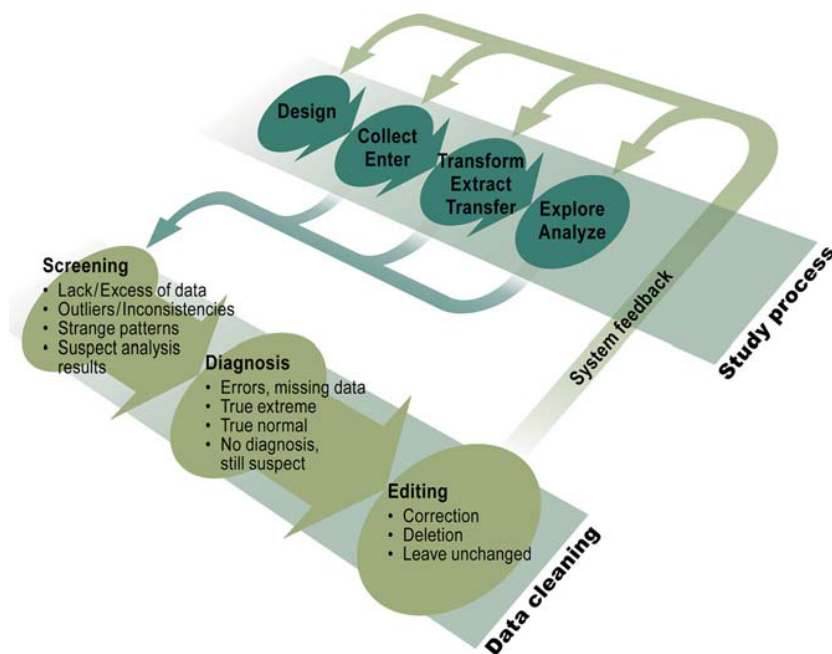


Figure 2.14: Framework for data cleaning (Van den Broeck *et al.*, 2005).

2.8 Repairable and non-repairable system reliability

When conducting reliability analysis, the definition of reliability allows for statistical and deterministic methods. Deterministic methods are largely focused around safety engineering and economic analysis. With the availability of statistical failure data, a hard scientific method exists for reliability calculation through observed occurrences. Further, it is important to determine whether the system under consideration can be modelled as repairable or non-repairable. Non-repairable components are discarded when they fail whereas repairable components can experience an improved condition. It is from the vantage point of repairable system reliability that literature is considered for the purposes of railway perway maintenance. This being said, there are certain types of repairable systems which are modelled with non-repairable theory due to certain unique system characteristics and so repairable system theory is loosely applied as well. There are two important definitions.

1) *Repairable system*: It is a system which, after failing to perform at least one of its required functions, can be restored to performing all of its required functions by any method, other than replacement of the entire system.

2) *Socket*: It is a circuit or equipment position which, at any given time,

holds a part of a given type.

It is important at this point to distinguish between the statistical analysis of a part and the analysis of system failure data. With parts, we are dealing with a distribution of time to a single failure whereas the times between successive failures of a system are modelled by a sequence of distribution functions, i.e., by a point process (Ascher and Feingold, 1984). Multiple parts must be tested to failure to generate a sequence of failure times, sufficient for statistical analysis. Failures of a single system are sufficient for statistical analysis if there are enough observed interarrival times. In an example of railway perway, a track section contains many parts, such as nuts, bolts, rail clips, sleepers, etc. and these parts make up a system (asset). The statistical failure of the section of perway can, therefore, be modelled by multiple failures from different parts in a reliability block network or multiple failures of a single system. The system approach is less data intensive and thus will be the focus of further investigation.

2.9 Two important reliability theory functions

There has been much ambiguity in literature when discussing the cornerstone principles of repairable systems as to what exactly has been defined. This ambiguity is clarified in this section. These functions are an important part of building a practical system reliability model.

2.9.1 Rate of occurrence of failures (ROCOF)

The rate of occurrence of failures function is the system point process model equivalent of the FOM. The ROCOF is an absolute rate. To understand ROCOF, we first have to consider a counting process $\{N(t), t \geq 0\}$. A counting process is a stochastic point process model that is described as a sequence of failures in time (a counting process), presented in Fig. 2.15. The expected value of $N(t)$ is denoted $V(t)$, $V(t) \equiv E[N(t)]$. $V(t)$ is assumed to be a continuous function. $v(t) \equiv V'(t)$ which is the time rate of change of an expected number of failures i.e. the rate of occurrence of failures (Ascher and Feingold, 1984).

2.9.2 Reliability

The reliability function is also known as the ‘survivor’ function. The reliability maps the probability that a part or system will survive past a certain time before failure. As reliability can be applied to parts or systems, the definition includes the potential for both interpretations. It is important that the reliability depends on the entire history of the process up to and including

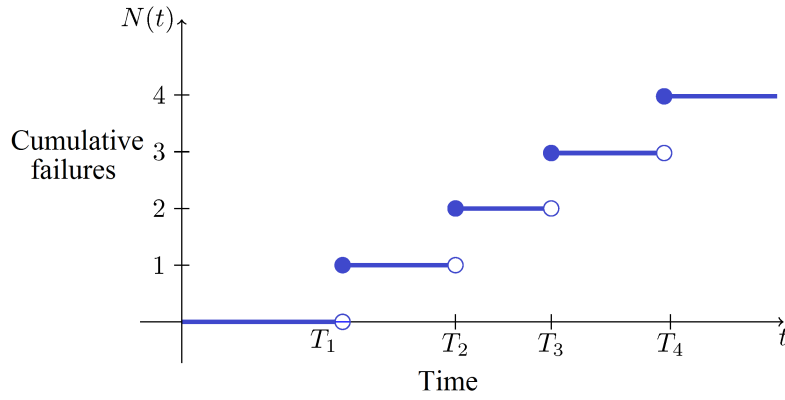


Figure 2.15: Counting process for a number of system failures.

the instant beginning the interval of interest. This condition is denoted $|H_t$ (Ascher and Feingold 1984, eqn. 2.7).

$$R(t, t + \tau) \equiv \Pr\{N(t, t + \tau) = 0 \mid H_t\} \quad (2.7)$$

2.10 Statistical analysis of system failure data

Statistical analysis of system failure data involves three stages during which the probabilistic model of the system is created, namely trend testing, parameter estimation and choosing the best fit for the selected point process model. These sections are discussed as well as scientific fundamentals to reduce the complexity encountered during analysis of such systems. In applying these three steps, we assume that data is available for each system considered and each system is analysed independently. The flow process detailing the steps in model design is presented in Fig. 2.16. From the figure, NHPP and renewal process blocks are of particular importance as these models are most commonly applied to system failure data. The NHPP model exists for data whose interarrivals are neither independent nor identically distributed due to presence of trends in the failure data. This is typical for repairable systems whose reliability characteristics are improving or deteriorating. If no trend is observed then data is identically distributed but not necessarily independent. Independence is often assumed when no trends are present as many data points are necessary to conduct independence tests. No trend results in a renewal process for which non-repairable system theory is applied. (Ascher and Feingold, 1984).

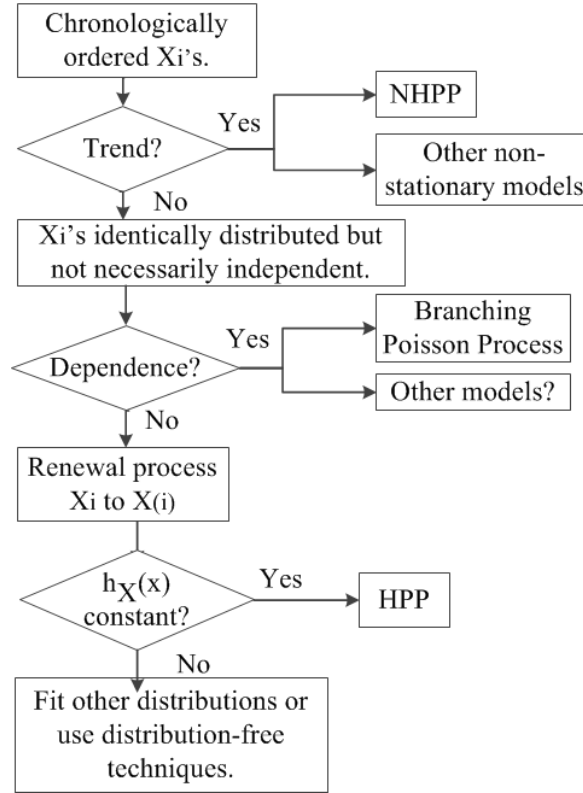


Figure 2.16: Statistical analysis decision tool for a point process (Ascher and Feingold, 1984).

2.11 Trend testing

Trend testing is concerned with the understanding of which point process model is applicable to the available failure data. The test weighs the statistical likelihood of a certain stationary sequence occurring, as the null hypothesis, opposed to a certain type of trend occurring, as the alternative hypothesis. With simplifying assumptions, this process helps categorise the failure sequence to a type of point process. A vast number of trend tests exist and a few of these are discussed. Additional tests such as the Military Handbook test (MIL-HDBK-189, 1981) and the Mann test (Loui *et al.*, 2009) can be understood from the literature.

2.11.1 Laplace

Null hypothesis - HPP | Alternative hypothesis - Monotonic trend

In a practical context, the number of failures in comparative systems varies quite drastically because each system is modelled independently. We discuss the specific situation where we have m pre-specified failures T_1, T_2, \dots, T_{m-1} .

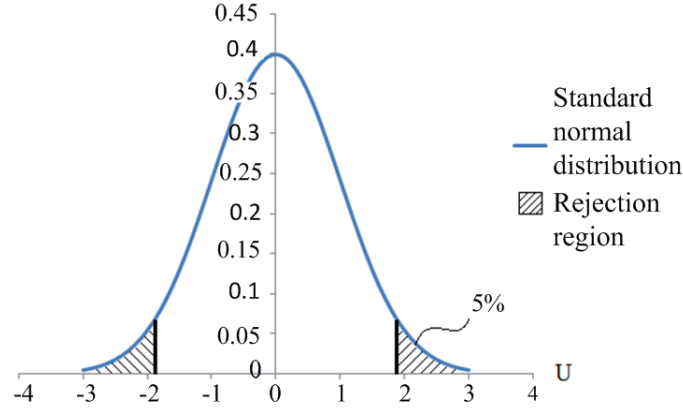


Figure 2.17: Standard normal distribution used in Laplace test.

Under the HPP assumption, the first m arrival times are the order statistics from a uniform distribution on $(0, T_m]$. This means that the probability of experiencing a medium length interarrival time as opposed to a long or short one is higher, as is the case for a uniform distribution. Equation 2.8 presents the uniform distribution, which approximates a standard normal variate with a 5% level of significance for $m \geq 4$ (Lawless 1982, eqn. 2.8).

$$U = \frac{\sum_{i=1}^{m-1} T_i}{\frac{m-1}{m-1} - \frac{T_m}{2}} \quad (2.8)$$

$$T_m \sqrt{\frac{1}{12(m-1)}}$$

$$z = \frac{\Upsilon - \mu}{\sigma} \quad (2.9)$$

Equation 2.8 is analogous to equation 2.9, which is the equation for a standardised score from a normal distribution; where Υ is the score of a unit in the population, μ is the mean of the population and σ is the standard deviation. This distribution is presented in Fig. 2.17. An example of this test is provided by considering successive numbers of inter-arrival times:

200, 195, 162, 145, 120, 104, 60, 49, 32, 2.

$U = +2.0$ for this test, which indicates a deteriorating system. Thus, if the U value falls within the 5% rejection region in this two-tailed test (< -1.96 or $> +1.96$) then it is conclusive that the point process does not fit a HPP and follows a monotonic trend, according to the alternative hypothesis. The Laplace test is optimum against the NHPP log linear model, which is presented here and referenced from a later section, eqn. (2.16). This test was recently

applied by Grobbelaar and Visser (2015) when determining a fit for component failures.

$$\rho_1(t) = e^{\alpha_0 + \alpha_1 t}$$

For small U values, it seems very unlikely that a trend exists for the data. Coetzee (2004) presented a case for the application of the Laplace trend test, stating that for low Laplace statistics, a renewal system is likely. He stated that any Laplace statistic within one standard deviation of the mean indicates that the data is non-committal as it closely follows a renewal process. A grey area therefore exists between one standard deviation and the 1.96 rejection value where no conclusions can be made about the data. Further trend tests are required in this case.

2.11.2 Lewis-Robinson

Null hypothesis - Renewal | Alternative hypothesis - Monotonic trend

The Lewis-Robinson test serves the same function as the Mann test (Loui *et al.*, 2009) but it does not use a ranking scale, rather a numerical scale is used which allows for the generation of a complete distribution (Wang and Coit, 2005). In this test, the interarrival times for the reverse arrangements are assumed to be independent under both the null and alternative hypothesis. This means that the nature of independence of the interarrivals is not known. The test statistic U_{LR} is formed (Lewis and Robinson 1974, eqn. 2.10) by modifying the Laplace test discussed above,

$$CV[X] = \frac{(Var[X])^{1/2}}{E[X]}$$

$$U_{LR} = \frac{U}{CV[X]} \quad (2.10)$$

Using the same inter-arrival times for the Laplace test example provided, $U_{LR} = +16$ which shows a large degradation past +1.96 that marks the start of the upper tier failure region for the test. The failure region for this test is applied in the same way as the Laplace test.

2.12 Point process models

A number of models have been applied to repairable systems; namely Homogeneous Poisson Process, Non-homogeneous Poisson Process, Renewal Process, Superimposed Renewal Process and Branching Poisson Process. These models are the basis for any failure prediction for systems. The powerful trend tests in literature are centred around the first of these three processes (Ascher and Feingold, 1984), with specific emphasis on Homogeneous and Non-homogeneous Poisson Processes.

2.12.1 Homogeneous Poisson Process (HPP)

The HPP is defined as a non-terminating sequence of independent and identically exponentially distributed X_i 's. Due to independence, each interarrival time has no knowledge of the previous and thus each failure follows an exponential probability of failure with time (Ascher and Feingold, 1984). The statement about identical distribution is the same as classifying stationarity. The mathematical definition of an HPP is as follows:

The counting process $\{N(t), t \geq 0\}$ qualifies as an HPP if

a) $N(0) = 0$

b) $\{N(t), t \geq 0\}$ has independent increments.

c) The number of failures in any interval of length $t_2 - t_1$ has a Poisson distribution with mean $\rho(t_2 - t_1)$. The Poisson formula with the mean substituted in is presented (Ascher and Feingold 1984, eqn. 2.11).

$$Pr\{N(t_2) - N(t_1) = j\} = \frac{e^{-\rho(t_2-t_1)}\{\rho(t_2-t_1)\}^j}{j!}, \quad (2.11)$$

for $j \geq 0$. ρ or in this case $v(t)$ is the constant rate of occurrence of failures (ROCOF). For the HPP, $v(t)$ is constant for both synchronous and asynchronous sampling. Knowing that the interarrivals are exponentially distributed, the definition of reliability and equation (2.11), the reliability function is

$$R(t_1, t_2) = e^{-\rho(t_2-t_1)} \quad (2.12)$$

2.12.2 Non-homogeneous Poisson process (NHPP)

The only difference between the NHPP and the HPP is that the rate of occurrence of failure varies with time for the NHPP. Although independent increments still exist, this means that the X_i 's of the NHPP are neither independent

nor identically distributed (Ascher and Feingold, 1984). The modification from the HPP to the NHPP is stated.

c) The number of failures in an interval of length $t_2 - t_1$ has a Poisson distribution with the mean $\int_{t_1}^{t_2} \rho(t)dt$, expressed mathematically as (Ascher and Feingold 1984, eqn. 2.13),

$$Pr\{N(t_2) - N(t_1) = j\} = \frac{e^{-\int_{t_1}^{t_2} \rho(t)dt} \{\int_{t_1}^{t_2} \rho(t)dt\}^j}{j!}, \quad (2.13)$$

for $j \geq 0$. From condition c), it follows that

$$E\{N(t_2 - t_1)\} = \int_{t_1}^{t_2} \rho(t)dt \quad (2.14)$$

From (2.13), the reliability function of the NHPP is

$$R(t_1, t_2) = e^{-\int_{t_1}^{t_2} \rho(t)dt} \quad (2.15)$$

The NHPP became a system model due to the fact that only a small percentage of a system's parts are replaced during repairs. Due to this fact, the assumption that the reliability before and after repair is approximately the same is reasonable. The two functional forms of the NHPP are the log linear model and the power law model. The **log linear** model (Lisnianski *et al.* 2010, eqn. 2.16) is expressed as

$$\rho_1(t) = e^{\alpha_0 + \alpha_1 t}, \quad -\infty < \alpha_0, \alpha_1 < \infty, t \geq 0 \quad (2.16)$$

Eqn. (2.16) is substituted into (2.14) generate the expected number of failures for the log linear model,

$$E\{N(t_2 - t_1)\} = \frac{e^{\alpha_0}}{\alpha_1} (e^{\alpha_1 t_2} - e^{\alpha_1 t_1}) \quad (2.17)$$

The form of the log linear model is substituted into equation (2.15) to determine the reliability,

$$R(t_1, t_2) = e^{-\frac{e^{\alpha_0}}{\alpha_1} (e^{\alpha_1 t_2} - e^{\alpha_1 t_1})} \quad (2.18)$$

The second form of the NHPP, the **power law** model (Crow 1990, eqn. 2.19) is presented.

$$\rho_2(t) = \lambda \beta t^{\beta-1}, \quad \lambda, \beta > 0, t \geq 0 \quad (2.19)$$

Considering the form of (2.16), the expected number of failures for the power law model is

$$E\{N(t_2 - t_1)\} = \lambda(t_2^\beta - t_1^\beta) \quad (2.20)$$

The reliability for this model can be easily computed as

$$R(t_1, t_2) = e^{-\lambda(t_2^\beta - t_1^\beta)} \quad (2.21)$$

2.12.3 IID data

After trend tests have been conducted on data and no trend is confirmed, a renewal process is typically assumed. If there is no evidence that the point process is not IID, then a homogeneous Poisson process is assumed. This assumption rests on a confirmed ‘no trend’ result from trend tests. Techniques to model this data are well known but fitting a Weibull distribution to the data is a technique which will specifically be considered here. Fitting a Weibull model to data is covered by Grobbelaar and Visser (2015). The inter-arrival times first need to be reordered in terms of magnitude, according to the definition of the cumulative distribution function

$$F_X(x) \equiv \Pr\{X \leq x\}$$

This means that if the data was not in order of magnitude, the probability estimates would be inaccurate due to issues of overlap, according to the definition. Parameters can then be estimated for the distribution for the Weibull cumulative distribution function (Al-Fawzan 2000, eqn. 2.22),

$$F_X(x) = 1 - e^{-(x/\eta)^\beta} \quad (2.22)$$

The researcher conducted a preliminary analysis on the case study and observed that the occurrence of ‘no trend’ data for which a Weibull fit is applicable would not be encountered. Simple analysis revealed that for the railway network under consideration, increasing trends would be dominant, although some ‘no trend’ data might exist. As a result, Weibull distributions were fit to data using parameter estimation methods discussed in future sections. In addition, the Chi-square goodness of fit test (Modarres *et al.*, 2009) and the Kolmogorov-Smirnov test (Sachs, 2013) were applied to the data. Through the application of these methods, the researcher is able to apply developed reliability models to a ‘no trend’ case. As ‘no trend’ cases weren’t encountered in the case study and ‘trend’ cases are expected to occur more frequently, reliability predictions and supporting methods for ‘no trend’ cases were not included in this document.

2.13 Parameter estimation

After trend tests are conducted on data to determine a point process model to be applied, the model needs to fit the data by estimating parameters as inputs into the selected point process equations. Different estimation techniques are

available, with their own advantages for application on different point process models. Of the many methods available, two of the most common are least squares estimation and maximum likelihood estimation which have been applied to datasets of small samples.

2.13.1 Maximum Likelihood Estimation (MLE)

To explain the maximum likelihood estimation procedure, let X_1, X_2, \dots, X_n denote n independent, identically distributed random variables with the density function $f(x; \theta_1, \theta_2, \dots, \theta_m)$, where f is of known form and $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ belongs to a subset Θ of m -dimensional space but is otherwise unknown. X_1, X_2, \dots, X_n could, for example, represent the lifetimes of n identical units in an inverse Gaussian distribution f . The joint density function of X_1, X_2, \dots, X_n is given by $\prod_{i=1}^n f(x_i; \theta)$ for fixed x_1, x_2, \dots, x_n . Say there is a fixed x_1, x_2, \dots, x_n for θ . This function is expressed as

$$\mathcal{L}(\theta; x_1, x_2, \dots, x_n) = \mathcal{L}(\theta; x) \quad (2.23)$$

$\mathcal{L}(\theta; x)$ is called the likelihood function. For a discrete distribution of X , $\mathcal{L}(\theta; x)$ is the probability of observing the values x_1, x_2, \dots, x_n for a given θ and thus indicates how likely it is to obtain the observations x_1, x_2, \dots, x_n for a given θ .

The first step in solving the MLE of θ , is to use the likelihood equation (Rausand and Høyland 2004, eqn. 2.24),

$$U_j = \frac{\partial \ln L(\theta; x)}{\partial \theta_j} = 0 \quad \text{for } j = 1, 2, \dots, m \quad (2.24)$$

Often non-linear equations are produced from a form of solution which can be processed using numerical methods (Rausand and Høyland, 2004). Note that the maximum likelihood method works for both parts and systems, thus the variable X can be replaced with the variable T .

2.13.2 Least Squares Estimation (LSE)

Least squares estimation follows the simple statistical curve fitting approach of plotting a line that produces the smallest square of the difference between expected and observed values. The parameter values that produce a least square error are the selected parameters. This can be expressed as

$$\min(\theta) = \sum_{i=1}^n [Y_i - f(x_i, \theta)]^2 \quad (2.25)$$

where $\theta = \theta_0, \theta_1, \dots, \theta_m$ which are the same parameters chosen for any point process model or distribution and Y_i is the observed value of a random variable.

Operations research methods are employed through the Excel solver tool to compute model parameters that produce the least squares against the observed failures from a dataset.

2.13.3 Non-homogenous Poisson Process

The two forms of the NHPP that were selected are discussed for their applicability in fitting model parameters. Lisnianski *et al.* (2010) used the maximum likelihood method to determine estimations for the **log linear** NHPP (Lisnianski *et al.* 2010, eqn. 2.26),

$$e^{\hat{\alpha}_0} = \frac{n\hat{\alpha}_1}{e^{\hat{\alpha}_1 T_n} - 1}, \quad (2.26)$$

$$\sum_{i=1}^n T_i = \frac{nT_n}{1 - e^{-\hat{\alpha}_1 T_n}} - \frac{n}{\hat{\alpha}_1}$$

The same MLE principle can be applied to determine the estimated parameters for the **power law** NHPP (Crow 1990, eqn. 2.27),

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}} \quad (2.27)$$

$$\hat{\lambda} = \frac{n}{T_n^{\hat{\beta}}}$$

These methods were applied to the arrival times of an exemplar system's failures,

190, 368, 511, 633, 745, 813, 864, 914, 946, 956

The Laplace trend test returned a statistic of +2.03, which indicates an increasing trend. For this, a NHPP log linear model was fit to the data using the MLE method such that $N(t) = 0.4595e^{0.0033t}$. The derivative of this function is the ROCOF form, $\rho_1(t) = e^{\alpha_0 + \alpha_1 t}$ where $\alpha_0 = -6.50$ and $\alpha_1 = 0.0033$. The LSE method was applied to the exponential function $N(t) = (e^{\alpha_0}/\alpha_1)e^{\alpha_1 t}$ to find the least square estimation of the log linear parameters. This produced $N(t) = 0.6482e^{0.0028t}$ with parameters $\alpha_0 = -6.31$ and $\alpha_1 = 0.0028$ for the ROCOF. Eqn. (2.25) is called the sum of square errors (SSE). The smallest SSE between the MLE and LSE is a basic indication of the best fit for the chosen model as the calculated fit most closely predicts the observed data in this case. The SSE for the MLE fit is 4.32, compared to the SSE of 0.873 for the LSE fit. The LSE more closely fits what actually happened to the system.

A NHPP power law model was fit to the failure data above for comparison to the log linear model. The MLE produced $N(t) = 7.32 \times 10^{-7} t^{2.39}$. The derivative of this function is the ROCOF form, $\rho_2(t) = \lambda \beta t^{\beta-1}$ where $\lambda = 7.32 \times 10^{-7}$ and $\beta = 2.39$. The LSE method was applied to the power function $N(t) = \lambda t^\beta$ to find the least square estimation of the power law parameters. This produced $N(t) = 1.31 \times 10^{-5} t^{1.96}$ with parameters $\lambda = 1.31 \times 10^{-5}$ and $\beta = 1.96$ for the ROCOF. The SSE for the MLE fit is 5.37, compared to the SSE of 2.70 for the LSE fit. The LSE more closely fits what actually happened to the system. In both cases, the LSE provided far better estimates of the actual data. The log linear model seems to provide a better fit in terms of SSE. A first estimate about model preference has been made with SSE. An argument is made for goodness of model fit in the following section.

2.14 Goodness of fit

Once models have been fit to failure data, the fit needs to be tested to ensure an adequate representation of the data. The Chi-square and Kolmogorov-Smirnov (K-S) tests exist to measure the observations from the data against the expected statistics from the model. These tests are applicable to Weibull distributions of IID data. The K-S test and Cramer von Mises tests do not apply easily to the power law and log linear NHPP models as they were not originally developed for the parametric case, according to Coetzee (1997). In addition, as NHPP models are point processes, a distribution exists for every inter-arrival time in the model. This means that goodness of fit tests such as K-S that test the fit of a single distribution are rendered ineffective. The Chi-square test can be applied however, although there normally aren't enough data points for an effective test. Alternative methods are considered as supporting selection of the best model for a NHPP.

2.14.1 Comparing log linear and power law models

In statistics, the goodness of fit of a model to data is calculated with the coefficient of determination (R^2). This coefficient explains the amount of error in a model, which is expressed as a performance percentage. The coefficient of determination is a robust explanation of fit as it accounts for variance of the mean of data as well as variance of each point prediction. These variances are explained by (Barrett 2000, eqn. 2.28)

$$R^2 = \frac{\sum (O_i - \bar{O})^2 - \sum (O_i - E_i)^2}{\sum (O_i - \bar{O})^2} \quad (2.28)$$

where the latent error of the model is explained by the first term in the numerator of eqn. (2.28). The second term in the numerator explains the error

between observations and predicted values. The coefficient of determination can be used to compare the fit of power law and log linear models, with the higher percentage indicating the better fit. Crowder *et al.* (1994) compare maximum likelihood values as a supplementary method to determine whether log linear or power law models offer a better fit. As least squares estimation often produces better parameters for the log linear and power law models, the maximum likelihood approach has limited application.

Crowder *et al.* (1994) linearised the log linear and power law models by applying the natural logarithm to the equations. They argue that the linearised model with the best fit straight line is a better model for the system. The form of the log linear equation is not appropriate for linearisation when it is expressed as cumulative number of failures against time. An estimate for the ROCOF of each equation is obtained by (Saldanha *et al.* 2001, eqn. 2.29)

$$v(\frac{1}{2}[T_{j-1} + T_j]) = \frac{N(T_j) - N(T_{j-1})}{T_j - T_{j-1}} \quad (2.29)$$

where the observation period $(0, T_m]$ is divided into m arbitrary intervals $(0, T_1], (T_1, T_2], \dots, (T_{k-1}, T_m]$. An approximate shape for the ROCOF is determined by the explained equation. Multiple iterations of interval selection should be conducted to ensure that the shape isn't attributed to the chosen subdivision. The linearisation method provides the same argument for best fit as the coefficient of determination as the same variances determine the best fit for each method. The R^2 method is preferred as it is less subjective than the linearisation method, but the linearisation method provides support from literature as it has been used for non-repairable reliability computations (Saldanha *et al.*, 2001).

2.14.2 Confidence bounds

As failure trend predictions are made from estimated ROCOF equations, one may be interested in knowing with what confidence predictions can be made. Confidence bounds are placed around estimates of observed data to state how much the actual trend can deviate from the model. Narrow confidence bounds at 95% confidence indicate that the observed data fits within the bands at 95% probability. Equations for confidence bounds on data are derived statistically. Cryer and Kellet (1986) shows that confidence for predictions decreases as time increases as the distant future is more uncertain than near future. This phenomenon explains why confidence bounds are very broad when predictions are being made from a time series. One would expect that statisticians predict with very sure confidence, but this is not the case in reality. Statistics simply uses the best methods available to make the best possible predictions. Vlok (2012) uses these arguments to place confidence bands around ROCOF

estimates. He first uses the Fisher information matrix to determine variance of power law or log linear parameters (Guo *et al.* 2010, eqn. 2.30),

$$\begin{bmatrix} Var(\hat{\theta}_0) & Cov(\hat{\theta}_0\hat{\theta}_1) \\ Cov(\hat{\theta}_0\hat{\theta}_1) & Var(\hat{\theta}_1) \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 I_\rho}{\hat{\theta}_0^2} & -\frac{\partial^2 I_\rho}{\hat{\theta}_0\hat{\theta}_1} \\ -\frac{\partial^2 I_\rho}{\hat{\theta}_0\hat{\theta}_1} & -\frac{\partial^2 I_\rho}{\hat{\theta}_1^2} \end{bmatrix}^{-1} \quad (2.30)$$

where $Var(\hat{\theta})$ is the variance of a estimated parameter, $Cov(\hat{\theta}_0\hat{\theta}_1)$ is the variance between two selected geometry parameters and I_ρ is the log likelihood of a selected ROCOF model. From the matrix, the variance of a ROCOF form is calculated (Vlok 2012, eqn. 2.31),

$$\begin{aligned} Var(\hat{\rho}(t)) &= \left(\frac{\partial \rho(t)}{\partial \hat{\theta}_0}\right)^2 \cdot Var(\hat{\theta}_0) + \left(\frac{\partial \rho(t)}{\partial \hat{\theta}_1}\right)^2 \cdot Var(\hat{\theta}_1) \dots \\ &+ 2 \cdot \left(\frac{\partial \rho(t)}{\partial \hat{\theta}_0}\right) \cdot \left(\frac{\partial \rho(t)}{\partial \hat{\theta}_1}\right) \cdot Cov(\hat{\theta}_0\hat{\theta}_1) \end{aligned} \quad (2.31)$$

The confidence bounds on a ROCOF model are finally determined using a statistical equation for confidence (Barrett 2000, eqn. 2.32),

$$\hat{\rho}(t) - z_\alpha \sqrt{Var(\hat{\rho}(t))} \leq \rho(t) \leq \hat{\rho}(t) + z_\alpha \sqrt{Var(\hat{\rho}(t))} \quad (2.32)$$

where $z_\alpha = 1.96$ for a 95% confidence bound. Confidence bounds can be compared for different models to determine which model has a tighter fit. It is desired to make reliability predictions with tight confidence to ensure accuracy of reliability estimates. In reality, confidence bounds on predictions in any time series are broad, but predictions are still made out of necessity. This is confirmed by Vlok (2012) who states that confidence on residual life estimates are broad. He further states that confidence can be significantly improved with the inclusion of condition maintenance data. If this is not available, the best predictions must be made with the data available.

2.15 Risk analysis

Risk has conventionally been used in engineering to calculate and mitigate hazards affecting people but risk has more recently been used as a maintenance prioritisation tool to save cost or improve reliability of systems. Risk is defined by Harnly (1998) as *Risk = Probability × Consequence*. This means that a risk is weighed according to the likelihood of it occurring and the severity of its occurrence. He further explains that risk assessment can either be qualitative or quantitative. Quantitative analyses require extensive statistical data, records and documentation whereas qualitative is more subjective. The idea of a risk matrix was formed here by which estimates for risk severity and

	I	II	III	IV	V
G	TNA	NA	NA	NA	NA
F	TNA	TNA	NA	NA	NA
E	TA	TNA	TNA	NA	NA
D	TA	TA	TNA	TNA	NA
C	A	TA	TA	TNA	TNA
B	A	A	TA	TA	TNA
A	A	A	A	TA	TA

HARD

	I	II	III	IV	V
G	TNA	TNA	NA	NA	NA
F	TA	TNA	TNA	NA	NA
E	TA	TA	TNA	TNA	NA
D	A	TA	TA	TNA	TNA
C	A	A	TA	TA	TNA
B	A	A	A	TA	TA
A	A	A	A	A	TA

STANDARD

	I	II	III	IV	V
G	TA	TNA	TNA	NA	NA
F	TA	TA	TNA	TNA	NA
E	A	TA	TA	TNA	TNA
D	A	A	TA	TA	TNA
C	A	A	A	TA	TA
B	A	A	A	A	TA
A	A	A	A	A	A

EASY

Figure 2.18: Risk matrix varieties (Markowski and Mannan, 2008).

occurrence are made. The risk matrix is considered to be the cornerstone of risk analysis as this tool remained throughout the development of risk analysis and it simply displays the results of rigorous methods.

2.15.1 Risk categories

Markowski and Mannan (2008) offer a 3D representation of a standard risk matrix and detail some important notions about risk matrices. They stated that prioritisation within a risk matrix has no external or empirical measurement outside of the risk matrix. This means that there exist multiple ways that high or low risk can be determined within a matrix. An example of three ways to arrange the risk categories within a matrix are presented in Fig. 2.18. The hard matrix is for a high cost maintenance scheme; the standard matrix represents a typical risk matrix in the process industry and the easy option is a low cost matrix, which also is less safe. The selection of risk categories will then depend on the desired maintenance strategy as well as real-case maintenance considerations such as cost and material on-hand. Markowski and Mannan (2008) used fuzzy numbers to quantify uncertainty in risk analysis to provide a more accurate 3D map of risk possibilities than for standard analysis. Fuzzification is only necessary for systems which are well understood, with plenty available data that have already been modelled as standard systems in the past.

2.15.2 Probabilistic methods

Khan *et al.* (2008) use a Markov (multi-state) model to calculate availability of components in a power plant, as the probabilistic arm of risk analysis. Sufficient statistical data was available to conduct a trend analysis which enabled availability calculations. In systems that have been running for a relatively short time, such as the tunnel lighting system investigated by Ng *et al.* (2003), probabilities need to be estimated qualitatively as data for statistical analy-

sis has not yet accumulated. Another way to calculate failure probabilities is the Bayesian approach, used by Paté-Cornell (2002). This method was used to identify human error in the installation and maintenance of heat shields in NASA aircraft. As large datasets were not available due to the small frequency of space missions, the expert reliant Bayesian approach was the most applicable way to evaluate likelihood of accident occurrence. He also used statistical data in another study to determine rate of accidents caused by anaesthetists in surgery. It appears that such methods are based on mean times between failures and are not probabilistic trend related. A unique approach to probability was suggested by Podofillini *et al.* (2006), with application to the railway industry using ultrasonic inspection of rail. They used the concept of the P-F intervals, which is the time between maintenance inspection of a fault and the failure caused by the fault. The average rate at which cracks propagate was tracked and a standard curve was drawn for probability of failure for a particular class of crack. This model can be viewed as a distribution describing failure of many identical parts. Trend analysis was not conducted on statistical data, but rather average failure rates with standard deviation were considered. These risk analysis options all have intuitive application to the case study at hand but incorporation of trend analysis of statistical data is lacking. This is perhaps a weakness of the proposed methods, that estimates of probability lack quantitative support.

2.15.3 Innovation in risk analysis

Risk analysis methods often evolve from specific needs and, for that reason, models seem individualistic. Khan *et al.* (2008) used monetary cost of failure and maintenance inspection as the severity in the risk equation (Khan *et al.* 2008, eqn. 2.33),

$$Risk = \sum_{i=1}^n c_i y_i^p \quad (2.33)$$

where y_i is the likelihood factor, expressed as availability. Another extrapolation of the classic risk equation is the severity index used by Harnly (1998), which is expressed as $SI = F \times (S + E)$. F is the failure potential, S is the safety factor and E is the economic impact on operations. This equation is useful in specific application but it is ill-advised to rank risk according to too many measures, but rather focus on the measure that is most important to minimise. A more appropriate individualistic risk analysis method is the probability versus cost approach by Podofillini *et al.* (2006) to minimise cost of railway track inspection while reducing the probability of derailment of trains. The power of this method is the detailed cost analysis incorporated and the route optimisation model to minimise cost of an inspection run. The method

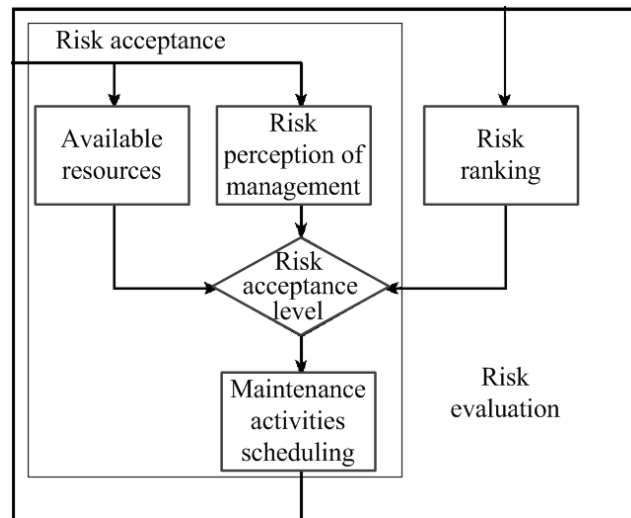


Figure 2.19: Risk in the business context (Ng *et al.*, 2003).

thus incorporates probability and severity in a focused manner to solve specifically observable and measurable problems.

2.15.4 Implementing a risk management programme

Ng *et al.* (2003) stress the importance of following a risk management process (RMP) to ensure that risks aren't simply categorised but that rankings from the risk analysis are implementable and well documented. The five core elements of risk analysis are: risk identification, risk measurement, risk assessment, risk evaluation, risk control and monitoring. This process is followed to not only prioritise risks but to transfer the analysis outputs to scheduled maintenance tasks. This process is presented in Fig. 2.19. They argue that resources, management support and risk acceptance on an operational level allows for the scheduling of maintenance. Harnly (1998) created a detailed procedure for risk management in the business context because maintenance inspectors were scheduling repairs and the operations department did not believe that correct priority was given to risks. His framework ensured that both inspectors and operations were consulted to arrive at risk rankings that were implementable. Backlund and Hannu (2002) investigated the translation from risk analysis to maintenance decisions and the roadblocks to effective implementation of strategy. They stressed that after risk estimation, there needs to be a paper trail whereby the risk analysis is verified and documented. If this is not done then maintenance cannot be prioritised as decision makers will not understand the steps followed to arrive at a qualified decision on risk ranking.

2.16 Summary of solution methods

The researcher analysed literature on solution methods to create a model that compares track sections for the sake of maintenance prioritisation. Solutions from many industries were sourced, including the passenger railway industry. It is forward thinking to identify solution methods even if they haven't been used for railway before. A research avenue is thus created in which problems can be tackled in more detail. Two potential solutions are discussed for further research consideration, such that practical implementation will arise.

2.16.1 Risk/reliability section model

For railway perway maintenance, it is important to know what the impact of a perway failure is, as well as how often failures occur. This information allows the maintenance engineer to be able to predict not only how often failures occur, but also the severity that a failure can be expected to have. The process can be described by a risk matrix model, where reliability indicates how likely failures will occur and average train delay caused by failures indicates the severity of the expected failures in a given track section. The reliability model can be applied to railway perway at three different levels: A track corridor, asset level and component level. The track corridor level is where failures are analysed at points anywhere between two train stations on a railway network. Asset level is where points along a track encompass a specific type of rail configuration, which is considered to be an asset and failures on an asset are tagged. Component level is where components that make up a perway asset are tracked for component failure.

For track corridor level, a criticality analysis needs to be conducted to understand the failure modes and their effect on the train network. The failure modes must be well understood as only failure modes that affect the operation of trains will affect perway reliability. This is according to the definition, as any component acting apart from design conditions (failure) is considered to be unreliable. Once the failure modes are understood, a dataset of track incidents must be obtained. This dataset is then filtered to only include failure modes that affect the operation of trains. A speed restriction or potential train stoppage is considered to be an effect worthy of being called unreliable. Repairable systems' modelling can be applied to the remaining data to determine reliability versus time curves. If no statistical trend is observed for data, then non-repairable system theory describes the system. This will identify track corridors that are most likely to experience the next failure. At asset level, the same procedure is followed to determine failure modes of track assets (which will be the same failure modes as for the corridors). In this model, data on each track asset is required to obtain reliability estimates for each asset. This model is set apart in that it shows maintenance personnel where to maintain

on a more specific level. The corridor level gives maintenance a map of where to maintain in the network whereas asset level gives a map of which assets to maintain in the corridor. For component level, a reliability block diagram can be developed for a track system based on reliability estimation methods for each track component. The basis of this reliability model is the probability of train derailment caused by geometric rail irregularities. Derailment is considered to be failure of the track system according to the definition, thus a reliability basis for the system is established. This method produces a single reliability statistic for each few meters of perway, from which a micro track maintenance schedule can be built. This model helps maintenance engineers to identify faults at component level, the smallest possible of the three levels.

Data capture needs to be more thorough, the lower the level of reliability analysis and a high failure rate over the entire network is necessary to capture enough failures consistently to generate an effective model. Failure analysis at the track corridor level is most likely to produce an effective analysis. This, coupled with estimates for failure mode severity, produces a risk metric for each track corridor, which can be compared between corridors.

2.16.2 Discussion of solution methods

The use of the existing TQI solution is an effective way of ensuring that track geometry is at a high standard. However, this method does not provide a scheduling solution for preventative work tasks and does not always eliminate the cause of poor track geometry. The reliability-based risk solution provides a comparative tool for preventative maintenance tasks, while mitigating root causes of failure. The TQI tool has some overlap with a reliability-based risk tool as the risk factor for track corridors is sometimes high as a result of the failures that have caused the deviation of track geometry. The overlap between these methods can be exploited to cross correlate the methods. The trend between TQI and reliability suggests that, for extreme cases of TQI, there should be a correlation with risk methods. Variables such as track age could prevent an establishment of the relationship between TQI and reliability. The failure to find such a relationship using this method does not necessarily mean that there is no relationship. The most practical comparison is made by comparing risk and TQI for a corridor with a high count of failure events.

Chapter 3

Case Study: PRASA Infrastructure Maintenance Priority Using Risk and Reliability

Every department in a passenger railway company has its own and unique challenges. At PRASA Infrastructure, fire-fighting tactics for maintenance are prevalent because many types of failures seem random and difficult to predict, according to an in-office civil engineer (personal communication, 9 March 2015). Condition-based maintenance and reactive maintenance is employed, using the IM2000 track recording car and maintenance teams respectively (Zaayman, 2011). There is value in condition-based maintenance as it allows for preventative maintenance, using a maintenance tool called track quality index (TQI). The shortfall of this method is that it has no direct link to failure modes of a perway and there is no direct link between TQI and failure prevention. It is anticipated that employing a scientific preventative maintenance method to railway infrastructure would allow for better control over maintenance activities (Song, 2009) and create a basis for failure prevention. This would prevent opportunity losses that are associated with reactive maintenance, caused by interruptions of service due to perway failure. Such a method would avoid the exorbitant cost of maintaining the railway perway to perfect condition, which is the only certain way to ensure perway reliability using condition-based methods.

For preventative maintenance to be cost effective, sections of perway need to be strategically maintained. This means that perway sections that are in danger of failure have a higher priority over healthy perway sections. It is important to be able to predict when perway failures will occur so that they can be prevented through foresight. Thus, an effect and criticality analysis as well as a fault tree analysis will be conducted on the PRASA tracks in the Western Cape Region to determine track failure modes that affect the operation of the train. If these failure modes can be predicted and prevented, then an effec-

tive preventative maintenance strategy can be created and the train service will become more reliable. Reliability methods are used for the prediction of failures, based on failure history to create a reliability model. This reliability model presents the likelihood of occurrence of perway failures. A risk metric then allows the prioritisation of perway sections for maintenance, based on the likelihood that the next failure will occur in a specific track section along with a failure severity. This distinguishes between track sections with low risk failures and track sections with high risk failures. A risk matrix is created using the reliability statistic and risk measure.

At PRASA, instead of condition-based maintenance aiding preventative maintenance, it seems that condition-based is in fact the primary means for maintenance. This is contrary to the findings of Golmakani and Pouresmaeeli (2014). Risk-based maintenance and preventative maintenance thresholds can be applied directly to the rail environment using FMECA and reliability centred maintenance. When a reliability threshold is decided, maintenance of a certain cost: performance ratio is established.

3.1 Maintenance investigation

Metrorail maintenance practices were investigated at the Salt River train depot, Cape Town to understand current maintenance performance and to identify improvement areas. This was done by interacting with personnel on-site as well as making observations. Literature on maintenance methods for multiple industrial applications is consulted to translate improvement methods to the railway environment.

3.1.1 Maintenance of the perway network

The Metrorail train network transports 1.7 million South Africans to work daily. The railway service is, therefore, an important driver of the South African economy. Fig. 3.1 presents a map of the network. Any reference to train stations is in association with this map. The perway division is responsible for maintaining track assets, ensuring that trains can pass safely from station to station without delay, obstruction or safety hazard. These topics were discussed in an infrastructure meeting that the researcher was invited to attend in May 2015. The systems manager from head office was presenting on maintenance frequencies for components based on statistical failure history. He was motivating maintenance staff to convert from corrective maintenance to a new preventative maintenance strategy. The staff was opposed to the idea as this required maintenance frequencies far beyond their capacity. Thus, it remains difficult to shift the percentage of corrective to preventative maintenance.

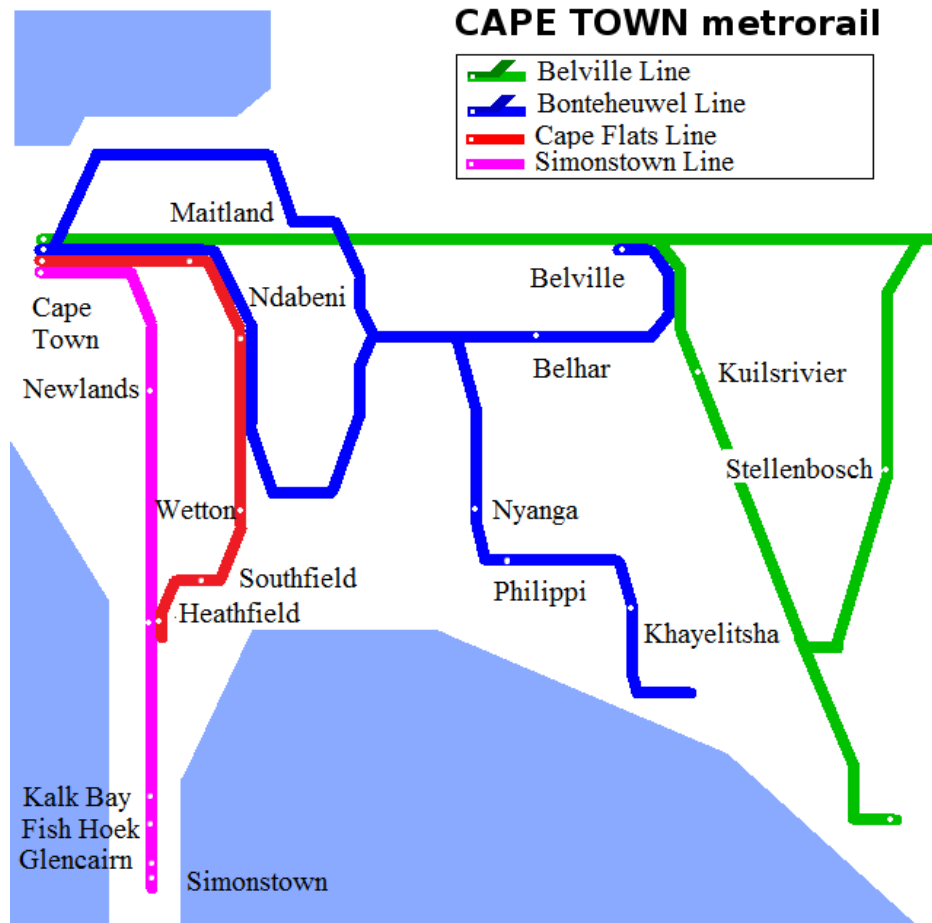


Figure 3.1: Reduced map of the Cape Town perway network.

nance in favour of more preventative maintenance. It is in this context that maintenance prioritisation will thrive.

3.1.2 Current maintenance performance

The pareto principle states that maintenance should be conducted 80/20 (80% preventative vs. 20% corrective). Based on PRASA's supplied perway failure history, an approximated maintenance representation is 50/50 (50% preventative vs. 50% corrective). This rule satisfied the need for a corrective and preventative maintenance strategy as discovered during problem case analysis. The only way to correct this maintenance imbalance is to increase maintenance productivity and conduct preventative maintenance more strategically. Other performance measurements were considered, which are adapted from operations performance indicators. Metrorail Cape Town has an availability of 78.8% for infrastructure as of 2014/15, which is short of the 95.9% target (professional communication, March 2015). This benchmark was considered as research of the problem case revealed its importance (Stenstrom *et al.*, 2013;

Line code	Train station A	Train station B	km marker 1	km marker 2	Asset name
OA	MAITLAND	WOLTEMADE	6.854	6.897	OA/TAN001
OA	MAITLAND	WOLTEMADE	6.337	6.914	OA/TAN046
OA	MAITLAND	WOLTEMADE	6.914	7.05	OA/CUR030
OA	MAITLAND	WOLTEMADE	6.897	7.066	OA/CUR001
OA	WOLTEMADE	MUTUAL	8.038	8.064	OA/TRN018
OA	WOLTEMADE	MUTUAL	8.03	8.132	OA/CUR007
OA	WOLTEMADE	MUTUAL	8.064	8.953	OA/TAN036
OA	WOLTEMADE	MUTUAL	8.132	9.132	OA/TAN009
OA	WOLTEMADE	MUTUAL	9.132	9.713	OA/TAN011

Figure 3.2: Database entry illustrating asset management levels for perway.

Åhrén, 2005). The delay minutes due to infrastructure faults has increased from 2013 to 2014. A Metrorail Cape Town infrastructure clerk (personal communication, 8 November 2014) reported 64 031 network delay minutes due to perway faults in 2013, which increased to 180 872 delay minutes in 2014.

3.1.3 Perway asset management practices

To keep track of perway assets, PRASA uses an incident list database populated by operations fault reports. In addition, a work order database which records call-outs on any maintenance job is used whether scheduled or unscheduled. Lastly, they use an asset register, which contains all perway assets and their location, according to a kilometre marker system. These three databases are consulted to understand what failures occur that affect the train service, what maintenance has been done on failed and functional perway assets as well what assets are currently part of the perway network. The perway network is managed at a line level, a system level and a sub-system level. The perway equivalent to these levels is: line code, sections between train stations and individual track assets. A simplified database entry is presented in Fig. 3.2 to illustrate this concept. The levels are analogous to the levels at which reliability analysis can be conducted. Scoping in to the right level ensures that failure data is available to coincide with the desired maintenance programme.

3.1.4 Meeting the standards of operations

The operations department presents the rolling stock and infrastructure departments with reliability standards to ensure that a specific level of performance is achieved by both departments. This allows for a buffer to manage operational uncertainty which allows operations to schedule the train service accordingly. The operations department expects a reliability of 90% for the train service, which means that both trains and track need to operate at a

reliability of 90%. Since reliability is time based, a future prediction of reliability starts from the current time, zero. This indicates that reliability will drop from a theoretical 100% at present time. The number of days of operation can be tracked from the current time until the reliability reaches 90% for a track section. This is an indication of the state of maintenance of perway as well as the ability of a perway corridor to pass a train safely. Reliability is theoretical but it is based on historical failures, thus, it can be viewed as feedback of failure prevention measures employed.

3.2 Track corridor reliability model

The developed reliability model is a series connection of perway assets that spans a track corridor between train stations. Between each track asset is a block joint asset which isolates the rail connections for the sake of a continuous electrical circuit from the train catenary, through the traction motors and into the ground through the rail. Block joints also exist between rail connections within a perway asset. The discussed configuration is presented in Fig. 3.3. A perway asset typically spans a few hundred metres or it can simply be a rail switch section for the trains to cross-over lines.

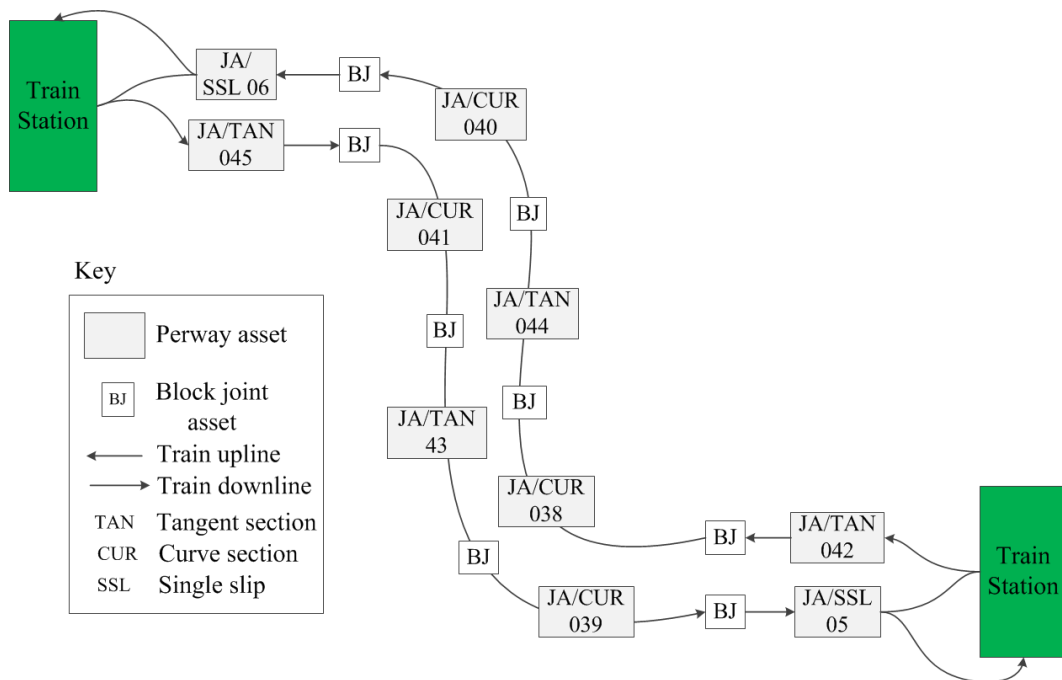


Figure 3.3: Reliability model of train network.

The current Metrorail track asset database spans 1227 tangent sections, 794 curve sections and 588 turnouts. Of the 588 turnouts; there are 313 one:

twelve turnouts, 215 one: nine turnouts, 28 single slips, 19 double slips, 8 diamond crossings and 5 scissor crossings. A small selection of these assets is considered for statistical analysis to prove applicability of the developed method. Within each perway asset is a grouping of components from the RBD shown in Fig. 2.3. Zooming in on this grouping is a means to break track tangent and curve assets into more detailed units. This zoomed in component level of reliability analysis is not further considered by the researcher due to the practical limitations imposed on data capture. It is sufficient to conduct reliability analysis for perway assets on a track corridor, even if the exact asset cannot be identified. The reliability on that corridor, after all, depends on the ability of the track to allow safe train passage according to operating conditions. If the failure mode of an asset is known, the line on which the failure occurred between stations and the time at which it occurred, system reliability analysis is possible for the track corridor.

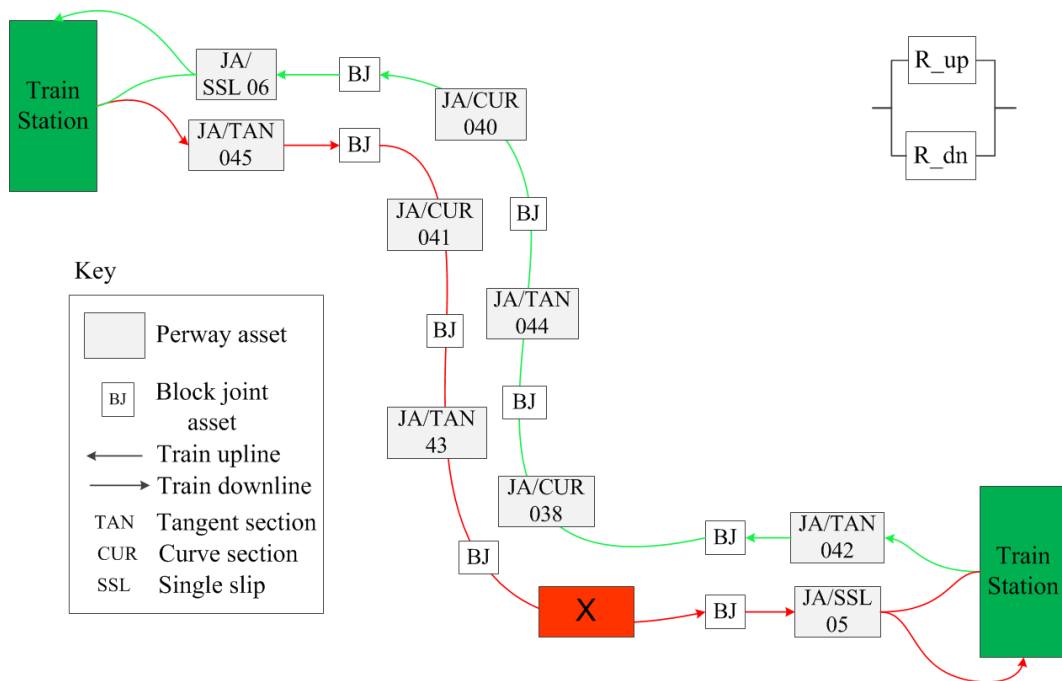


Figure 3.4: Reliability model of train network.

One would expect that a train can travel in only one direction from station A to station B and any failure that affects the train operation would be added to a database of asset failures on that section. This assumption would result in an overly simplified reliability analysis for world class railway networks. In practical operation for international railway companies, when a track asset on a line fails and the train cannot pass, the train is often switched to the line where traffic flows in the opposite direction. The extended model is presented in Fig. 3.4. The traffic switching sometimes causes little or no delay but during

peak times, there is a more significant amount of delay. This extra option adds redundancy into a reliability analysis. For Metrorail, traffic problems prevent switching of trains and so Fig. 3.3 is the reliability model of choice.

3.3 Failure mode identification and criticality analysis

As discussed in the section 2.4, the FMECA methodology was applied to PRASA perway. An interview was conducted at the PRASA train depot in Salt River, with the Western Cape's Southern and Northern line track inspector. From the interview, track failure modes were identified as well as the amount of corrective time necessary to make repairs. He was consulted to determine the ease of detection of the failure modes using standard inspection methods. From the operations department, a data log of track failures was collected to correlate the failure modes with average train delay time. This enabled the determination of the severity of each failure mode. Failure occurrence figures were obtained from the incident data sheet for perway in the Western Cape, supplied by PRASA head office. The RPN tables for occurrence, severity and detection evaluation are presented in Tables A.1, A.2 and A.3 respectively.

Multiple failure modes were considered from both the incident and work order databases. These databases pertain to any incident affecting the operation of trains and maintenance work conducted on the track at PRASA. Incidents with failure modes that have negligible severity in terms of train delay or disruption were omitted from the list. A final list of perway failure modes was created and is presented in Table 3.1. Note that for entries with the same RPN, the detection value is omitted to determine priority for ranking. This RPN ranking of failure modes is a first attempt to understand the critical perway failures of the Western Cape passenger rail network.

3.4 Root cause analysis

Perway failure modes from the FMECA were investigated for probable causes of failure and tree diagrams were constructed for each failure mode. A list of root causes for each failure mode was created from the FTA, as a collection of key problems to be solved by track maintenance. This list of root causes provides a means to identify maintenance tasks which prevent critical perway failure modes. An example of a fault tree for rail breaks is provided in Fig. 3.5. From the identified root causes, current maintenance tasks were better understood and new maintenance tasks were identified that are of preventative nature. These tasks can be applied in conjunction with a preventative maintenance strategy, which identifies problems before they become critical

Table 3.1: RPN table for railway perway failure modes (Chin *et al.*, 2009), (Xu *et al.*, 2002).

Failure mode	Occurrence	Detection	Severity	RPN
Missing rail clips	9	8	8	576
Sand on track	7	8	8	448
Twist	9	7	7	441
Cracks	4	9	7	252
Faulty block joint	9	7	4	252
Skid marks	6	8	5	240
Broken crossing (frog)	9	3	8	216
Horizontal alignment	4	7	7	196
Broken rail crown	5	5	7	175
Dirty ballast	6	7	4	168
Wide/narrow gauge	7	3	7	147
Vertical alignment	3	7	7	147
Rail breaks	8	2	8	128
Slack in rail	8	4	4	128
Environmental damage	5	3	8	120
Broken crossing blade	7	2	8	112
Pantograph hook-up	7	2	7	98
Points not closing	7	2	7	98
Super elevation	2	7	7	98
Kick-out	6	2	8	96
Corrugation	4	3	8	96
Broken sleeper	6	4	4	96
Derailment	5	2	9	90
Side wear	4	3	7	84
Obstruction	6	2	6	72
Rust	4	2	7	56
Points overlap	6	3	3	54

and, thus, create a platform for prevention through maintenance. Valuable current maintenance practises can be used in accordance with preventative maintenance to ensure that the most efficient and cost effective maintenance is conducted. This root cause analysis study proves that knowing what to maintain due to insight into perway reliability has a practical outworking into maintenance tasks. This allows a maintenance engineer to maintain unreliable track sections as there are possible maintenance practices that he/she can employ to bring track sections back to a reliable state.

The researcher argues here that the problem with the current track maintenance is not only that the best maintenance tasks are not being done, but also not being done at the optimal time while the priority for maintenance tasks remains unclear. The lack of priority leads to lack of vision with regard to effective maintenance. The result of this blurred vision is corrective maintenance, which is an intensive maintenance strategy. To avoid this, the researcher suggests reliability priority for track sections. If the maintenance teams apply this to practise, certain maintenance tasks will continue as normal but some tasks will be done on a priority basis, ensuring that maintenance is conducted on unreliable track sections before reliable ones. Ultimately, the cost will increase as thorough maintenance is conducted from track section to track section but these will decrease when less sudden failures result because maintenance occurs at the right place, at the right time. The service life of assets will increase, thus, increasing asset utilisation, which reduces cost. Table 3.2 was constructed by evaluating root causes and critical paths from the FTA. From these, the researcher suggests possible preventative maintenance tasks for the top ten failure modes in terms of risk. These tasks are proof that a maintenance priority strategy is possible to execute. The fault trees for these failure modes are presented in appendix B.

3.5 Risk matrix

A risk matrix is created to answer the question: ‘Which perway failures are going to have the highest impact on the operation of the train service and where are these failures most likely to occur?’ This question needs to be answered in order to improve maintenance prioritisation of perway. The power of the risk matrix is its statistical probability of survival of a perway corridor and average train delay caused by a failure historically. This concept is similar to the probability severity model created by Kumar *et al.* (2010), as discussed in section 2.2.1. The probability aspect of this risk matrix predicts the chance of a failure occurring on the track section, which increases as time progresses from last failure. The average historical delay provides an estimate of severity for a potential failure. With this information, one can rank perway corridors according to a risk metric, which in effect, stipulates the likelihood and severity

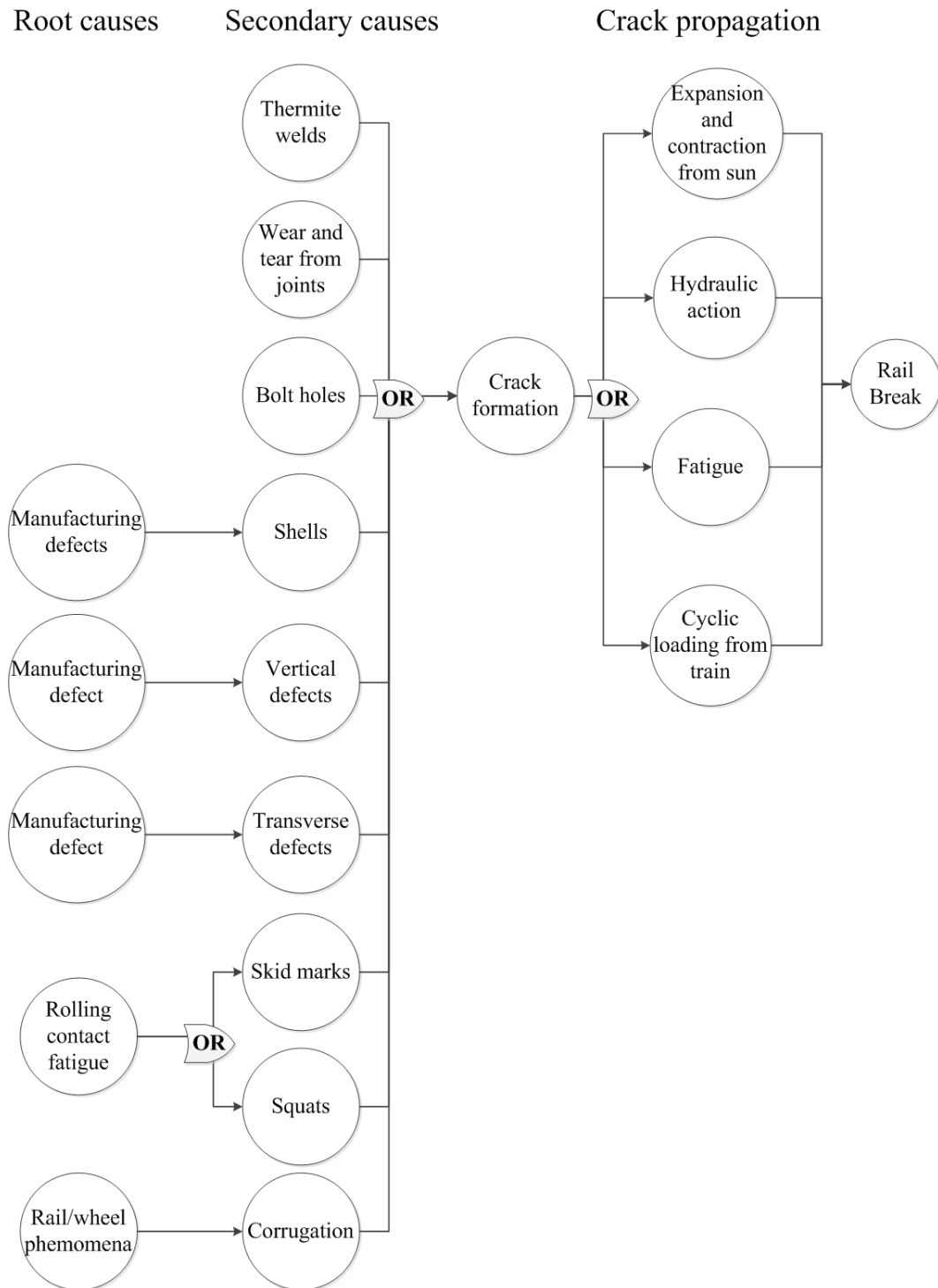


Figure 3.5: Fault tree for rail break failure mode.

Table 3.2: Identified preventative maintenance actions for ten specified failure modes.

Failure mode	Possible preventative maintenance action
Missing rail clips	Regular inspection, replace clips before reports, anti-theft installations.
Sand on track	Wind breakers at hot spots, regular clearing.
Twist	Regular inspection, sleeper checks, ballast checks, tamping and packing.
Cracks	Regular inspection, grinding, welding, quality check incoming rails, coat rails.
Faulty block joint	Regular cleaning, regular inspection, install bolts, remove cracked units.
Skid marks	Quality check incoming rails, observe during foot patrols.
Broken crossing	Regular inspection, replace parts before reports, check for cracks.
Horizontal alignment	Regular inspection, sleeper checks, ballast checks, tamping and packing.
Broken rail crown	Regular grinding, correct geometry defects, replace rail before critical.
Dirty ballast	Remove and replace ballast, tamp ballast, repack subgrade.

of failures occurring in that corridor. If the pareto principle is indeed correct, then 20% of all the track corridors will be causing 80% of the delay and the risk matrix makes the identification of these corridors possible. The discussed matrix is presented in Fig. 3.6.

The probability of survival discussed above is the reliability of a perway corridor. This means that the likelihood arm (vertical scale) of the risk matrix is specified by the reliability performance of a perway corridor. The higher the reliability of the perway corridor the lower the likelihood of failure gets. Thus, the reliability performance of a perway corridor is interpreted as the predicted number of days from the last failure until the reliability has degraded to 90%. This 90% is the expected perway achievement by the operations department at PRASA. So, the reliability performance is measured by the amount of time, within which the system is able to maintain a satisfactory reliability level for

1/2 +	10	20	30	40	50	60	70	80	90	100
1/4 +	5	10	15	20	25	30	35	40	45	50
1/6 +	3.33	6.67	10	13.3	16.7	20	23.3	26.7	30	33.3
1/8 +	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
1/10 +	2	4	6	8	10	12	14	16	18	20
1/12 +	1.67	3.33	5	6.67	8.33	10	11.7	13.3	15	16.7
1/14 +	1.43	2.86	4.29	5.71	7.14	8.57	10	11.4	12.9	14.3
1/16 +	1.25	2.5	3.75	5	6.25	7.5	8.75	10	11.3	12.5
1/18 +	1.11	2.22	3.33	4.44	5.56	6.67	7.78	8.89	10	11.1
1/20 +	1	2	3	4	5	6	7	8	9	10
	0+	20+	40+	60+	80+	100+	120+	140+	160+	180+
Expected delay per incident for a track corridor (mins)										

16+	Tamping, replacements, PM, inspection across entire corridor, planned maintenance.
8.5+	Tamping, replacements of critical failure modes, PM of critical failure modes, planned maintenance.
4+	General inspection, PM and replacement of faults found.
0+	General inspection, do maintenance as previously.

Figure 3.6: Risk matrix model for perway maintenance prioritisation.

operations. The lower the number of days, the faster the system degrades to the 90% level. This degradation is expressed as a factor, 1/days so that the risk level increases with an increasing number of 1/days (as the likelihood increases, so must the risk). The 1/days scale is sensitive and so the range can be adjusted if required. On the horizontal scale of the risk matrix is the average delay minutes per incident for a track corridor. The average network delay per failure mode is calculated by taking the delay per incident from operations reports and averaging them for a time period. When analysing the failure statistics for a track corridor, the average delay per failure mode can be applied to each type of failure in the record and, thus, an expected delay is calculated. A short range is applied to the expected delay in the risk matrix as the values represent averages and, therefore, would not differ significantly from each other. The values entered into the exemplar matrix provide an estimate of what a populated matrix would look like. The risk categories in the matrix are defined according to Markowski and Mannan (2008), in Fig. 2.18. The hard matrix style is preferred as the top level categories in an easy matrix would be unpopulated in this risk analysis as the majority of data is expected to fall in the middle and lower tiers of the matrix. Fig. 3.6 describes an exemplar maintenance plan for each matrix category. The more critical the risk, the more extensive the maintenance, when focus is on maintenance that makes measurable improvement on the reliability of a perway corridor. An example calculation is provided to explain how the matrix entries are obtained.

From Nov 2008 - Nov 2014, the track corridor between Kalk Bay and Fish Hoek experienced failures due to consecutive missing rail clips and faulty block joints. An average network delay for each of the two failure modes was calculated from operations reports and adjusted by information received by the Metrorail Cape Town track inspector. There were eight counts of rail clip faults and each count caused an average delay of 150 minutes in addition to four counts of block joint faults, of which each caused an average delay of 20 minutes. The average corridor delay is then $(8 \times 150 + 4 \times 20)/12 = 106.7$ minutes. Statistical analysis indicates that it takes the corridor 31 days to degrade to 90% reliability, thus, from the risk matrix, the risk factor is 3.4 min/day. This means that with every passing day, a theoretical 3.4 minutes are added on, assuming that a failure occurs as the perway reliability standard drops below the required operating condition of 90%. An explanation of risk categories is provided in chapter 6.

Chapter 4

Data Analysis

Data analysis involves the cleaning of raw data and input into developed models to produce the desired output. The developed reliability model and severity calculations for train delay are tested by evaluating actual failure data and determining statistical trends from it. The purpose of the data analysis chapter is to demonstrate that theory on failure statistics can be applied to the case study in order to calculate reliability for perway corridors. Further, it is important that the severity of failure modes is realistic, which is confirmed through analysis. A single perway corridor is mostly studied for the purposes of detailed discussion.

4.1 Data sets

A thorough data set study was conducted on the available maintenance failure data at PRASA Infrastructure to understand how maintenance and failures are tracked. Three databases were considered; namely, the events log, work order database and perway asset register from November 2008 to January 2014. The assets in the register are current as of the year 2015. The objective of this data study was to create a sortable and clean data-set for all maintenance events, which is a combination of the clean events log and work orders databases. From the clean data, failure trends could be predicted for the rest of 2014, into 2015. The events log is a database for perway failures that affected the operation of trains. This resulted in track maintenance personnel conducting work such that regular service was restored. This list gives a good indication of what failures caused an interruption in the train service. The work order database is important as it contains maintenance information for the condition-based restoration of track geometry parameters, which are indicators for perway failure modes that will not be recognised by operations. Dirty ballast and any track replacements are identified through the work order database. The track renewals are important when the events and work order databases are combined as a change in an observed reliability trend could have

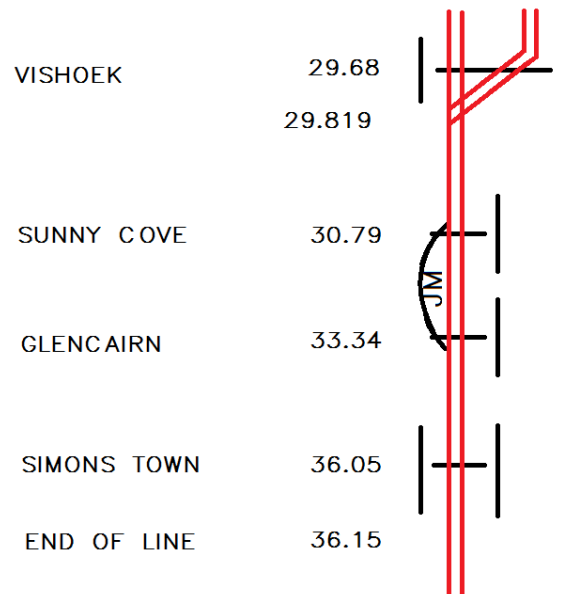


Figure 4.1: Line map extract of the Metrorail perway network.

been caused by a major event. A major event could be extensive renewals, rendering the data for reliability analysis unusable up until the change. When a perway failure event occurs, the train station that is geographically closest to the failure is logged as the location of the failure in the events log. In addition, a kilometre marker is recorded, which is the equivalent of co-ordinates in the perway domain. An asset tag is assigned to the failure, in accordance with the assets specified in the asset register.

The asset register was designed to trace every section of perway such that the location of a perway failure is identified. Each asset is described by a line code, an asset type and an asset number for the asset type. The asset register additionally has kilometre markings for each asset. The kilometre marking system only works in accordance with a line code that specifies the major line on which the asset is located. For example, KM 10.400 exists on the Wetton - Ottery track corridor as well as the Netreg - Heideveld track corridor. The difference between them is that the Wetton - Ottery corridor is on the Cape Flats line, which has line code JN whereas the Netreg - Heideveld corridor has line code MD. The limitation of the asset register is that it is not immediately obvious between which train stations the asset is found. Consequently, the researcher consulted a line map and updated the asset register by linking the kilometre markings and line codes from the asset register to the line map presented in Fig. 4.1. Only an extract of the line map is presented here.

As not every perway event is logged with complete data, a secondary

Action date	WO description	Asset tag	Work order no.	Keyword	Event or WO
2014/04/02	CLOVELLY BROKEN RAIL KM 28/508H CAM	JL/CUR061	14-001797	BROKEN RAIL	Event
2014/07/23	CLOVELLY KM 28/747 BROKEN JOINT PLA	JL/CUR063	14-004161	BLOCK JOINT	Event
2014/07/30	KALKBAY - CLOVALLY SLACK AND BROKEN	JL/CUR059	14-004253	SLACK	Event
2014/08/15	FISH HOEK BROKEN CROSSING IN 3rd RD KJ	JL/TRN025	14-004639	CROSSING	Event
2014/09/10	Repair gauge on curve /Repair twist/Lift and tamp	JL/CUR058	14-014375-000 -		WO
2014/09/10	Replace broken sleepers and fastenings	JL/CUR058	14-014375-000 -		WO
2014/09/10	Repair twist lift and tamp joints and box in ballast	JL/CUR063	14-014376-000 -		WO
2014/09/23	KALKBAY SLACK AT KM 28/508 DN LINEPAN	1M	14-005447	SLACK	Event
2014/09/26	CLOVELLY GAUGE TO WIDE AND SLEEPERS	1M	14-005503	WIDE GAUGE	Event

Figure 4.2: Combined work order and events database extract between Fish Hoek and Kalk Bay.

database (asset register) was manipulated to unlock hidden information in the primary database (events log) for the recovery of omitted necessary information. This data manipulation is an example of data cleaning where there is missing data, as shown in Fig. 2.14. In this instance, the missing data can be retrieved. Fig. 4.2 provides an extract from the combined events and maintenance data-set along the Fish Hoek - Kalk Bay track corridor. An asset tag of 1M is observed in the extract. After investigation, it was uncovered that instead of assigning an asset tag during the work order by the maintenance personnel, a general tag was assigned to this asset. This general tag sometimes makes it impossible to distinguish which one of two track corridors the asset is in. Assumptions have to be made for the placement of this asset. To clean the database to the observed state, duplicate data entries had to be removed. Many description cells contained multiple entries in one which were split by a cell to column algorithm. IF and FIND Excel function algorithms were created to sort the descriptions into keywords which represent failure modes. This is how failure mode entries were tracked.

4.2 Reliability methods applied

The entire Metrorail perway network is one large system of interconnected tracks, all working to serve for effective public transport for the greater Cape Town area. The perway failures in the system as a whole were observed to better understand what characteristics occur in the many perway corridors that make-up the network at large. Fig. 4.3 presents the number of perway failures per year for the entire network, from January 2008 - November 2014. The failures increase from 268 per year in 2008 to 545 in 2014. The failure rate in 2014 is more than double of that in 2008. From eyeball analysis, it seems that the system as a whole observes an increasing trend; meaning that the reliability of the network is deteriorating. This is most likely due to maintenance that is not preventively focused. It seems that maintenance is not improving the reliability of systems indicating that perway corridors are more likely to experience an increasing trend as opposed to a decreasing one.

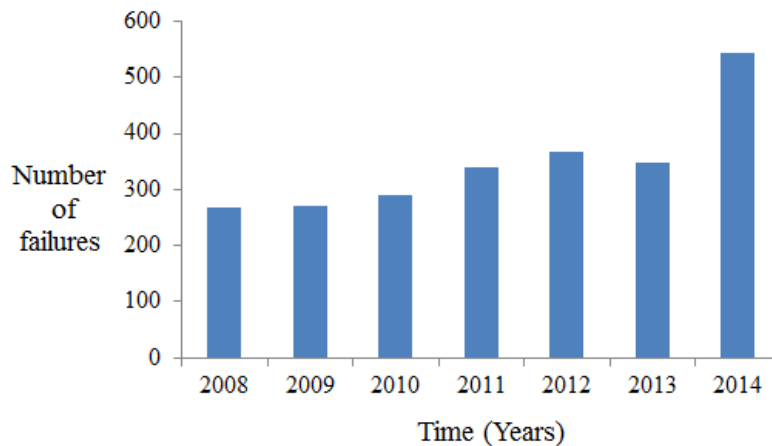


Figure 4.3: Number of perway failures for Metrorail Cape Town.

A non-committal trend is also likely but not as likely as an increasing one. The cumulative number of failures for the network was plotted against time for the period discussed. This type of graph is a good trend indicator and is a standard tool used when fitting a model to failure data, although this method is applied here purely as a speculative tool. The Laplace test was applied to the data and a statistic of 13.04 resulted, which indicates a heavily increasing trend. The power law and log linear models were fit to the data to determine which was the best fit. Using the power law, there is a 66% chance that a failure will occur within the first day of operation of the system after its last recorded failure. Although some error may factor into the large scale of this analysis, an initial understanding of the vastness of Metrorail's perway network has been conceptualised. One could argue that this chance of failure used to be less, by taking a train ride from Stellenbosch to Cape Town. A passenger worries that the journey takes almost twice as long as it did five years ago and ponders what has changed that the experience has worsened.

4.2.1 Analysis of a single perway corridor

Repairable system theory can be applied to any perway section that has four or more failures and no full shed maintenance upsets for that corridor. Maintenance has an effect on reliability which is difficult to quantify, meaning that any model for reliability includes the effects of historical maintenance. With more failures present, trends in data become obvious because precision increases. Thus, a track corridor was sought that had a higher number of failures. High traffic routes such as Cape Town - Woodstock and Belville - Tygerberg, as shown in Fig. 3.1, were not considered for analysis. This is because a reasonably small percentage of recorded failures for each track corridor were assigned to a general asset tag and the location for these assets was not determinable

for these cases. Now, to trace a failure to a track corridor, it must first be determined between which train stations the failure took place, then on which coded line the failure took place and, finally, whether the failure occurred on the up traffic or down traffic direction of track within the coded line. For a corridor such as Cape Town - Woodstock, there are four coded lines, each with an up line and down line. Thus, any perway failure between these points could have occurred on one of eight sections of track. For such data, estimating the location of failure is next to impossible for busy track corridors and errors are inevitable in this set up. As this error cannot be ignored, sections where smaller errors are present are preferred for analysis. The Simonstown - Glencairn track corridor was selected due to its poor perway condition and single multi-directional traffic line.

The researcher acquired TQI data for the Metrorail perway network and it was discovered that the Fish Hoek - Simonstown track corridor, consisting of four train stations, had the worst recorded TQI of any track corridor between January 2009 and December 2013. This 2.71km corridor measured an average TQI of 19.5, which is far above the 7.5 threshold for mainline tracks. The locations of failures for this corridor were identified using a decision support diagram presented in Fig. 4.4. This track corridor had the most recorded perway failures of any of the nearby corridors, which sparked the interest of the researcher. Understanding the reliability trends of such a perway corridor would shed some light to unveil the characteristics of tracks that deviate far from their design conditions. The events and work order database entries were scrutinised for all assets within the Simonstown - Glencairn perway corridor. Two duplicate entries were removed as well as any other perway failures resulting from obstructions and environmental effects such as sand on tracks. No component theft was experienced on this track corridor. A few track replacements were recorded, as well as welding and some localised tamping on tracks. Very little TQI improvement was recorded at any stage during the 2009 - 2013 period, which confirms that only localised maintenance was conducted. The NHPP model is perfect for this scenario as it assumes that the reliability of the corridor will remain largely intact when such localised maintenance is conducted, which is much the same as the car's reliability when you change a tyre. On the 5th of May 2011, a gauge repair was conducted due to extremely poor gauge measurements on track asset JM/CUR026. This failure event, although slightly delayed in its recording, was added to the data set. Train drivers generally report heavy vibrations caused by such track geometric deviations and then repair work is logged in the work orders database.

The cleaned data set for the Simonstown - Glencairn perway corridor was modelled as a set of arrival times of a repairable system. The corridor is represented by a string of assets each in respective sockets, as discussed in section 2.8. As this perway corridor only has a single line between the two stations, its

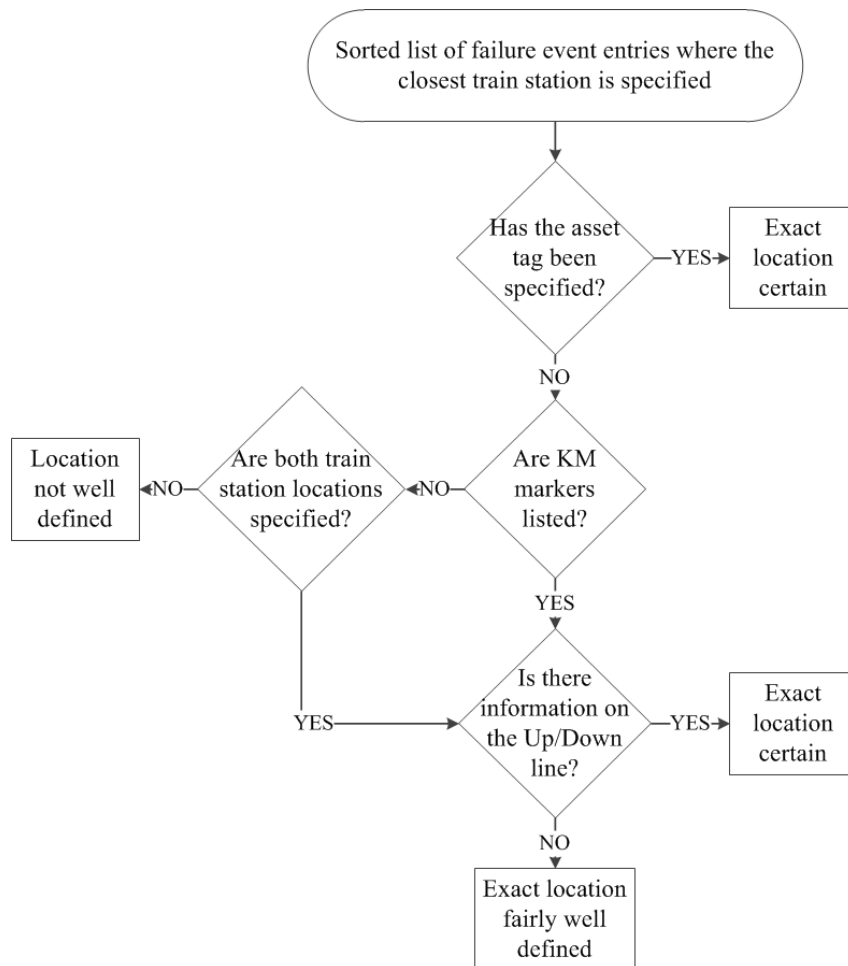


Figure 4.4: Decision support diagram for the location identification of perway failure events.

reliability is represented by one of the reliability model arms, presented in Fig. 3.3. For this model, unpredictable failures resulting from theft, environment and obstruction were excluded because they are not caused by track degradation and, thus, would skew the model. The arrival times of perway failures are as follows:

249, 501, 1158, 1697, 1979, 2084, 2089, 2101, 2133, 2167, 2175, 2187, 2285, 2328, 2420 (days).

The data was left truncated to remove any uncertainty that arises when the last failure event before the start of recording is unknown. As there was plenty of data available for trend estimation, time zero was recorded from 3 February 2008 instead of 2 January 2008 when failure data recording began. The data was right truncated as well, as the last failure occurred on the 17th of October 2014, close to 1 November 2014 when the last available data was collected. Any

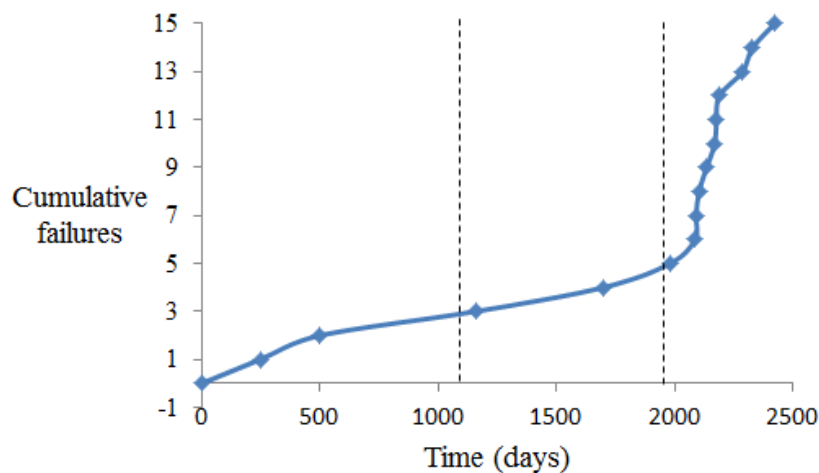


Figure 4.5: Cumulative failure against time for the Simonstown - Glencairn perway corridor.

trends were thus plotted from failure to failure, similar to a case when suspensions are assumed at the beginning and end of data-sets to simplify analysis. A cumulative failure against time plot was constructed to visualise the failure trend of the data, presented in Fig. 4.5. For the first 1200 days of analysis, a decreasing trend is prevalent for the data. The trend gradually changes around this time, where failures occurred more and more rapidly. At around time 2000, the increase experiences a rapid jump. The initial jump eventually eases off slightly. It is evident that the dataset exhibits three distinct periods each with a different trend. Nothing extraordinary happened at time 2000 in terms of maintenance that could have upset the trend, according to the work order database. The general trend for the entire time history resembles the effect of deformation of an object when it encounters a load. Such an object seems to be healthy under initial small elastic deformation, but under heavier loads, clearly begins to deteriorate. As plastic deformation of the object is experienced, the object exponentially deteriorates. If the perway failure trend loosely follows track condition, a critical condition must have been reached around time 2000. Effective preventative perway maintenance is important, to avoid such deterioration. PRASA currently uses a MTBF to account for reliability in their perway network. For this corridor, the estimated MTBF is 202 days. The error of comparing this value with other perway corridors for maintenance planning is that the MTBF ignores the trend that is observed from the data. The MTBF is likely to underestimate the failure rate for the data in this case. A more intelligent analysis is necessary to predict the reliability of the corridor.

A Laplace statistic of +3.13 was calculated from the Laplace trend test, which indicates a heavily increasing trend. This result is in accordance with

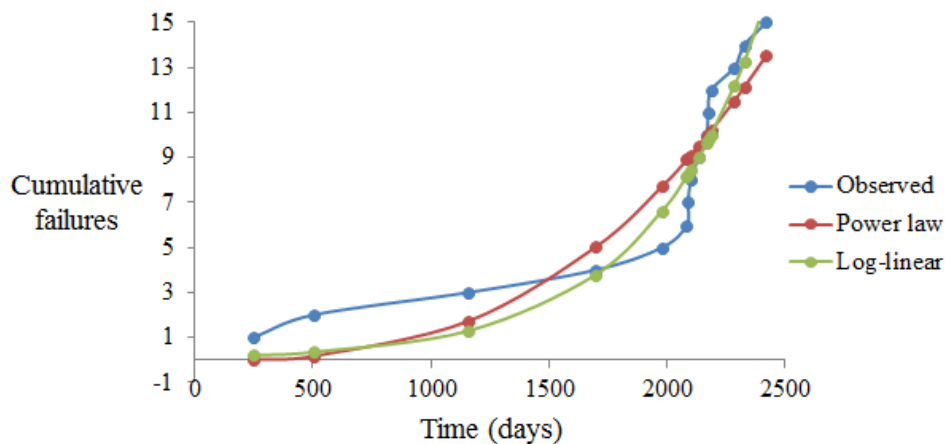


Figure 4.6: Log-linear and power law model fit to failure data from the Simonstown - Glencairn perway corridor.

eyeball analysis and fits the general trend for perway corridors for the entire Metrorail network. Fig. 2.16 indicates that an NHPP model can be fit to this data. A power law or log linear model is suitable for fit, although the best of the fits will model the data to ensure goodness of fit. The MLE method and the LSE methods were tested on multiple data sets but the MLE estimates seemed crude compared to LSE estimates, thus only the LSE was continued. The log linear least squares estimation produced parameters $\alpha_0 = -8.26902$ and $\alpha_1 = 0.001992$. The R^2 value for these parameters is 92.4%. The power law least squares estimation produced parameters $\lambda = 4.669 \times 10^{-9}$ and $\beta = 2.7963$, with an R^2 value of 89.8%. The log linear model seems to represent what actually happened to the system better than the power law. The graphical fits are presented in Fig. 4.6. It is clear from the graphic that the expected values from log linear model are closer to the observed values than the power law fit, especially for $N = 4$ and up. Until at least $t = 2000$, the power law model doesn't accurately fit the observed failures which makes it an undesirable choice when compared to the log linear model. Goodness of fit for four other perway corridors is presented in appendix C.

A linear estimation method by Crowder *et al.* (1994) is applied to both log linear and power law models to confirm R^2 results. Fig. 4.7 presents the linearised ROCOF form for log linear and power law models. The estimates in time were made for the mean time of each selected estimated interval, which explains each point on the respective graphs. This method provides the same argument as the coefficient of determination, which shows that literature supports R^2 comparisons. The log linear model is clearly the most linear of the two models. Linear estimation for four other perway corridors is presented in appendix D.

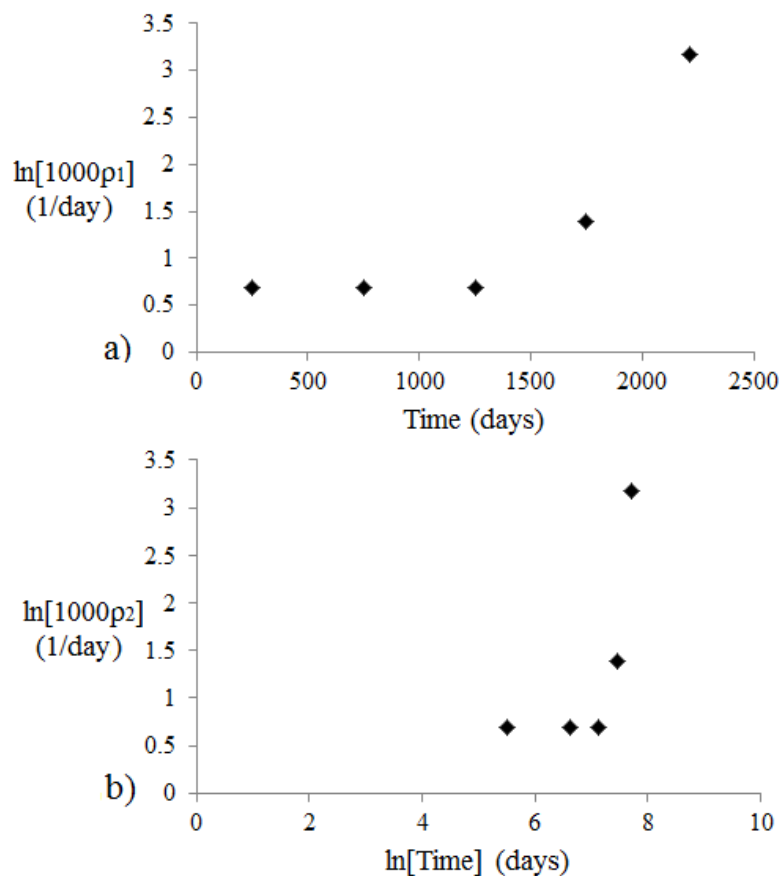


Figure 4.7: Estimated ROCOF for a) log linear and b) power law models.

Confidence bounds were placed around the ROCOF estimates for both log linear and power law models, ρ_1 and ρ_2 respectively. The Fisher information matrix allows the calculation of these bounds, which provides confidence for a model fit and future predictions (section 2.14.1). These bands give an indication of the likelihood that the actual ROCOF, which has been estimated, falls inside the bound region. The 95% confidence bounds for log linear and power law models are presented in Fig. 4.8. The power law model has tighter bounds at the end of the observation period which indicates that predictions from the power law model can be made with more accuracy. Although this is true, it would be erroneous to conclude that the power law is the better model from this comparison as the confidence does not include the goodness of model fit based on regression. Although the power law model provides more confidence, the log linear model is chosen as it is a more accurate model of the observed data. Confidence bands were applied to the ROCOF of four other perway corridors in appendix E.

Section 3.5 discusses that the operations target for perway reliability is 90%.

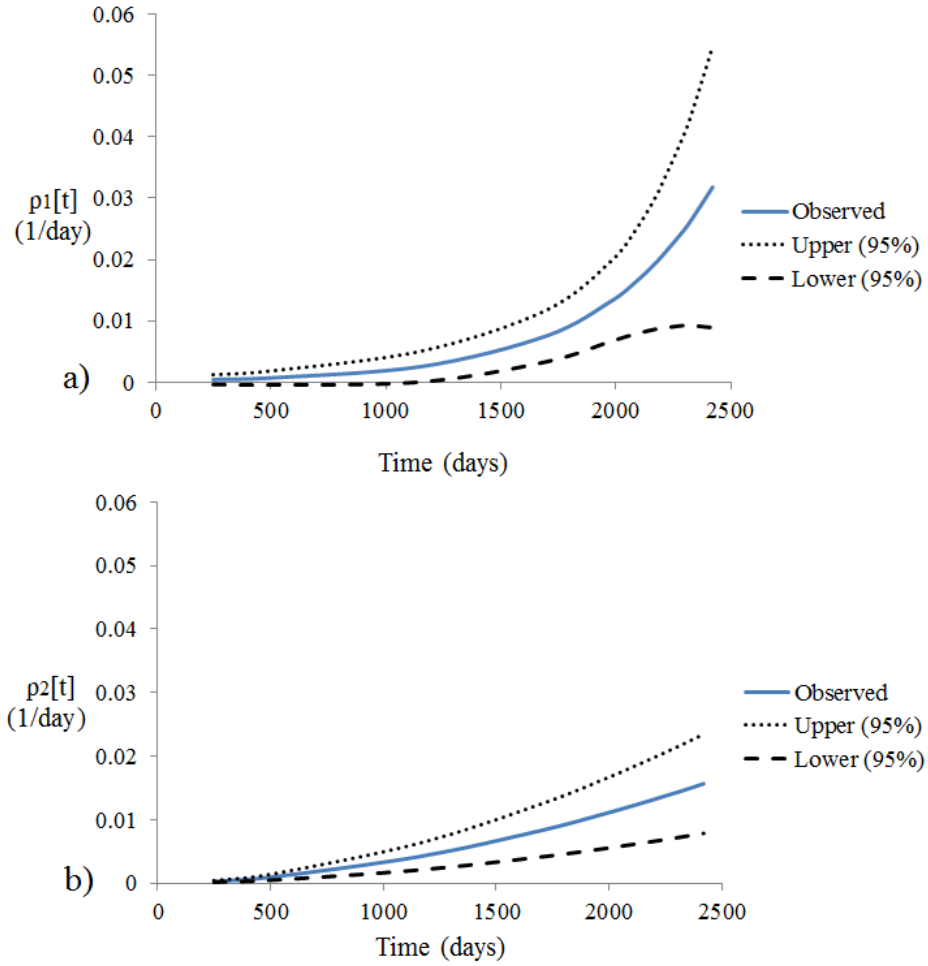


Figure 4.8: ROCOF with 95% upper and lower confidence bounds for a) log linear and b) power law models.

Equations 2.18 and 2.21 calculate the reliability over the interval t_1 to t_2 . The model exists to predict the future reliability, and so reliability is calculated from the time of last failure, $t_1 = 2420$ onwards. The reliability against time curves are plotted for the log linear and power law fits for the first 150 days after the last failure. These curves are presented in Fig. 4.9. The reliability predictions for both fits are close enough to each other to provide confirmation that the reliability is representative of the data. The reliability from the log linear fit falls more steeply than the power law model. 90% reliability is experienced after 3.3 days of operation for the log linear fit and after 6.7 days for the power law model. As the log linear model provides the best fit, it is confirmed that a reliability of 90+% can be maintained for only three days after the last failure. A poor reliability was expected as this perway corridor has the worst condition of any other in the perway network, according to TQI. This low reliability is justified in the following section. Operations cannot effectively

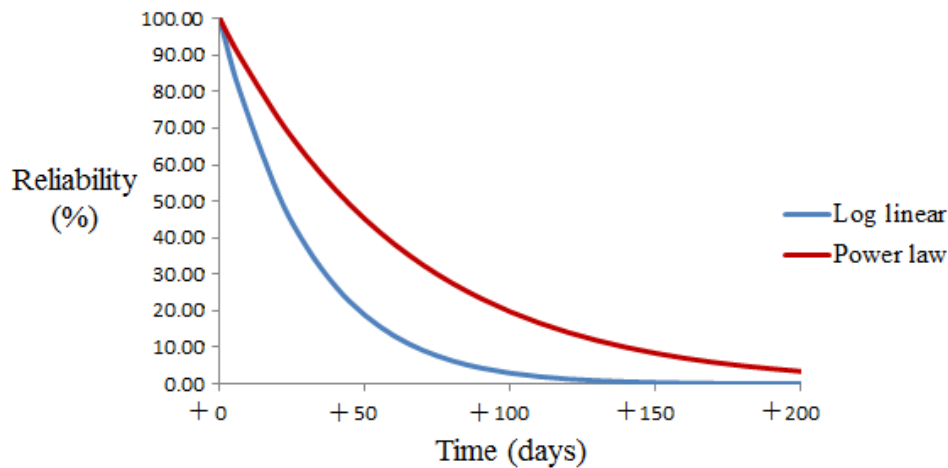


Figure 4.9: Reliability of power law and log linear models for the perway failures of the Simonstown - Glencairn corridor.

schedule for this perway corridor as the reliability is so poor. Extensive renewal is recommended.

4.2.2 Scrutiny of system reliability

Keeping perway at a high reliability through maintenance is a complex challenge, considering that perway covers a large geographical area. It is difficult to track the changes in reliability due to maintenance when sparse maintenance is conducted. Reliability tracking is only possible if maintenance is conducted on all sets of failure mode specific components for which an estimated increase in reliability is calculated. It is further difficult to build redundancy into the system as infrastructure requires large capital expenditure and redundancy adds significantly to that cost. The current perway network has some redundancy at component level, by adding in more sleepers and rail clips than necessary for operation. Rails lack redundancy, which is why they are one of the primary causes of risk in a railway network. The only way to add rail redundancy is to build additional track sections in the direction of train travel, which is not cost efficient. The current state of railway infrastructure is a single direction perway corridor, which is the only possible traffic route. This corridor is a series connection of perway assets. If any one of these assets fail, train traffic ceases until the asset is restored. The Simonstown - Glencairn corridor has 18 perway assets on its main line. Assuming that each asset was brand new and the reliability of each asset degrades evenly, the following trend presented in Fig. 4.10 would occur. The system reliability decreases exponentially as the reliability of each asset decreases. This means that the reliability of each asset needs to be maintained at a high standard to avoid rapid deterioration at a system level. If all the assets have a reliability of 95%, then the system reli-

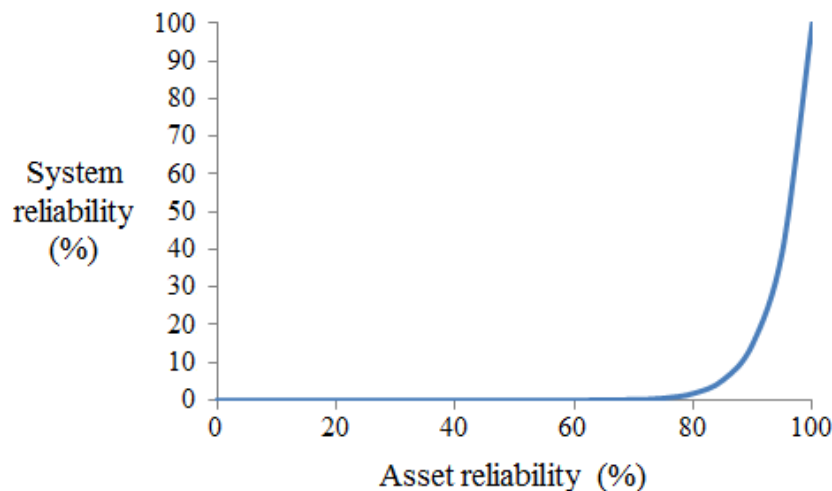


Figure 4.10: Theoretical degradation of system reliability with asset reliability for Simonstown - Glencairn perway corridor.

bility is $0.95^{18} = 40\%$ at a moment in time. If all the assets except one have a reliability of 95% and that single asset has a reliability of 10%, the system reliability drops to 4%. The difference between 40% and 4% is determined by just one asset. In reality, new assets and old ones would be mixed together at any given time due to replacements. Thankfully, because the service life of railway components is generally 20 years, system reliability is manageable for shorter time horizons. For a poor condition perway corridor such as Simonstown - Glencairn, the state of reliability is understood by the system against asset trend presented.

4.3 Pareto analysis of perway failure mode severity

The Pareto principle states that 80% of effects come from 20% of causes. This principle is applied to severity of perway failures, which is understood through the train delay they cause to the railway network for the period of January 2008 to November 2014. More severe perway failures require immediate track closure while less severe failures allow the train service to continue with an applied speed restriction. Different network delay results from each of these instances. The Pareto diagram of Fig. 4.11 presents the estimated network delay time for each failure mode based on the delay minutes caused by each in operations records.

The top six events in the Pareto analysis are supposed to cause 80% of delay time. The total recorded delay time over this period, according to Metrorail,

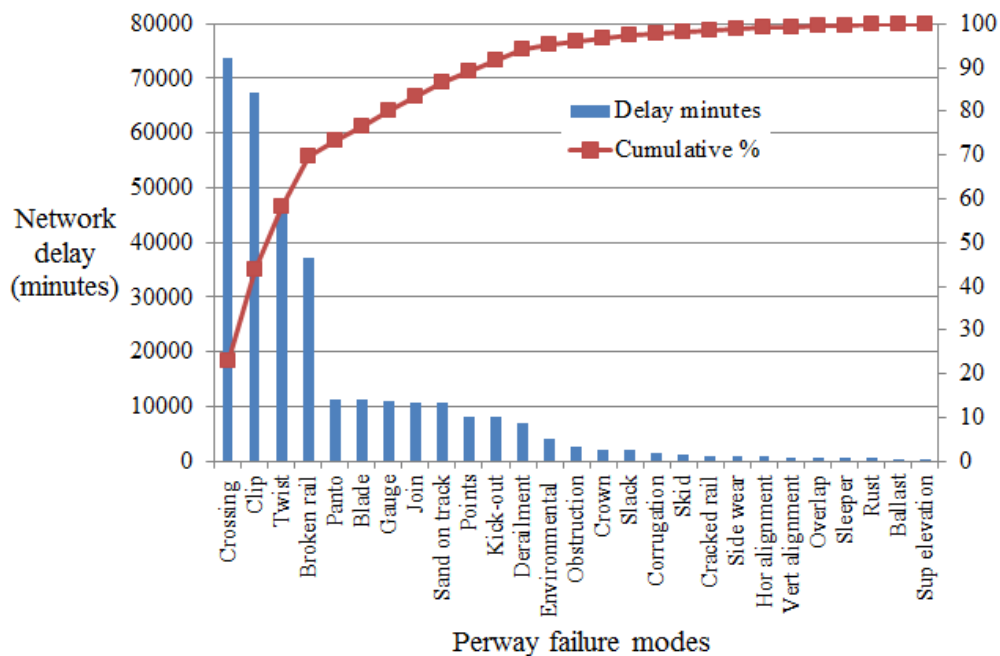


Figure 4.11: Pareto diagram of train delay caused by each perway failure mode.

is approximately 350 000 minutes. The researcher tracked 322 449 delay minutes, which is not a far-off recorded value. The deviation is the result of a sample being taken from operations reports to calculate the average delay of each failure mode, which has a small deviation from the population of failures. Metrorail has a clerk who records total failure minutes, which is what the researcher's calculations are compared to. The top six failure modes cause 77% of the total failure time, which is in accordance with the Pareto principle. The reasons for the high delay for the top six failure modes are explained. Careful attention is paid to these failure modes as they are principle causers of perway unreliability and also possibly deterioration of track.

Crossing delay minutes are high because they account for 320/2524 recorded events, which produced an occurrence value of 9 towards its RPN. This, in addition with the 230 minutes average delay expected from such an event gives it a high delay ranking. Its high delay results from the urgency of repair for such an event. If there is rail damage at a crossing, train delay risk is high as this is where a train changes from one set of tracks to another. Grinding and welding is often needed for such a repair, which adds waiting time until a repair can commence. The rail clip failures are high as they account for 450/2524 events. Normally, a rail clip failure wouldn't be extremely critical but these recorded failures are a result of theft, for which 50+ clips can be stolen at a time. With such a large number of consecutive missing clips, the rail can easily kick out; causing large stresses in the rail, rail breaks and an unsuitable

channel for a train to travel on. The 150 minutes of average delay are a result of immediate track closure and because such faults are often only found in the morning during peak traffic, after theft from the night before. Twist repairs result from maintenance which has been planned, specifically when twist geometric track parameters exceed their operational limit. 303 of these events were recorded with 154 minutes expected delay for each. Recordings of twist and gauge repairs dominate repairs over other known geometric parameters. The researcher believes that the other track geometry maintenance events were omitted as they are often repaired during twist maintenance. For these kinds of failure modes, speed limits are initially placed on the perway corridor when failure is identified. The next maintenance step is occupation of the perway, which is out of peak traffic. Delay here is caused by the occupation taking longer than scheduled and due to the reduction of network capacity due to the line shut off. There are portions of the network with only a single line to pass traffic. An occupation here causes large disruptions.

Broken rails account for 218/2524 events for which a delay is rated at 180 minutes. Rail breaks cause an immediate disruption of the train service and no trains can pass over this section. A new rail has to be fetched from the equipment stores and a welder is needed to help fasten the new rail onto the section. As this is one of the most severe types of perway failure according to logic, this justifies the appearance of this failure mode in the top six. A train pantograph hook-up usually takes a few hours to rewire. This is not directly a perway fault but it requires occupation of perway to fix. 80/2524 events recorded were pantograph hook-ups with an average delay of 141 minutes. Trains obviously can't pass while the electrics are down thus immediate maintenance needs to be conducted for this failure mode. Blade repairs are related to crossing repairs as they are the tip of the new connecting rail at a switch and need to be intact to avoid train derailment. Blade repairs account for 55/2524 events and have an average delay of 203 minutes. The high delay is explained by the sudden discovery of the fault which leads to immediate occupation, taking a few hours as there is grinding and welding work necessary. The top severity failure modes are justified and this builds confidence for the application of severity data to determine risk.

4.4 Risk matrix applied

The severity ranking for the Simonstown - Glencairn perway corridor was calculated using the average delay per failure, explained in section 3.5 and the number of days until 90% reliability. The average delay per failure is presented in Table 4.1 for the corridor. Notice that there are common occurring failures with small delay implications. If these were the only failures present then one would conclude that if the efficiency of minor fixes is improved, then

maintenance problems would be solved for this corridor. This is not the case here because multiple rail breaks are also present, as well as a derailment, which has extremely severe consequences for network reliability. In addition, this perway corridor deviates from design conditions with wide gauge problems. The poor reliability of the corridor is worsened by the diversity of failure modes experienced and by eyeball analysis, which seems to be shifting into a high risk category.

Table 4.1: Failure modes and their average delay experienced on the Simonstown - Glencairn perway corridor.

Arrival time (days)	Failure mode	Average delay (mins)
0	Block joint	19
249	Block joint	19
501	Broken rail	170
1158	Wide gauge	159
1697	Derailment	374
1979	Broken rail	170
2084	Block joint	19
2089	Block joint	19
2101	Broken rail	170
2133	Block joint	19
2167	Block joint	19
2175	Block joint	19
2187	Block joint	19
2285	Slack	15
2328	Block joint	19
2420	Slack	15

For the corridor, the average delay is 77.8 minutes and the reliability metric 3.3 days. That results in 23.6 minutes/day. The placement of the risk factor on the risk matrix is presented in Fig. 4.12. The corridor fits into the highest risk maintenance category, which is desired. The corridor is unreliable and it is in a bad condition. The maintenance decision that results from the risk matrix is a complete overhaul on the corridor, which is the conclusion that TQI and reliability information points to. As severe failures are part of the unreliability of the track, the high risk status of this corridor is solidified. The reason for the corridor risk not falling deep within the risk category, considering its high risk is because it is not desired that many perway corridors fall into this category. A complete perway overhaul costs millions of Rand and only

1/2 +	10	20	30	40	50	60	70	80	90	100
1/4 +	5	10	15	20 X	25	30	35	40	45	50
1/6 +	3.33	6.67	10	13.3	16.7	20	23.3	26.7	30	33.3
1/8 +	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
1/10 +	2	4	6	8	10	12	14	16	18	20
1/12 +	1.67	3.33	5	6.67	8.33	10	11.7	13.3	15	16.7
1/14 +	1.43	2.86	4.29	5.71	7.14	8.57	10	11.4	12.9	14.3
1/16 +	1.25	2.5	3.75	5	6.25	7.5	8.75	10	11.3	12.5
1/18 +	1.11	2.22	3.33	4.44	5.56	6.67	7.78	8.89	10	11.1
1/20 +	1	2	3	4	5	6	7	8	9	10
	0+	20+	40+	60+	80+	100+	120+	140+	160+	180+

Expected delay per incident for a track corridor (mins)

16+	Tamping, replacements, PM, <u>inspection across entire corridor, planned maintenance.</u>
8.5+	Tamping, replacements of critical failure modes, <u>PM of critical failure modes, planned maintenance.</u>
4+	General inspection, PM and replacement of faults found.
0+	General inspection, do maintenance as previously.

Figure 4.12: Risk matrix for the Simonstown - Glencairn perway corridor.

very few of the perway corridors in the network can fall into this category. The analysis conclusion is that the reliability-based risk method facilitates the prioritisation of perway corridors for maintenance. The validation of this method is presented in the following section.

Chapter 5

Results and Validation

The reliability analysis of a single perway corridor was extended to show that there is consistency between reliability predictions when multiple perway corridors are compared. The analysis of these corridors is the pinnacle of the research and presents the results of the topic. It is not enough to simply apply the developed risk model to a real case, but validation of the model is necessary to justify its continued use. Data from the condition tool TQI was correlated to the developed risk comparison tool to show that extreme risk cases do relate to extreme states of perway condition.

5.1 Reliability analysis of multiple perway corridors

The failure data for five perway corridors was assessed for a recorded history between January 2008 and November 2014. The names of the corridors are: Simonstown - Glencairn, Kalk Bay - Fish Hoek (up), Philippi - Nyanga (up), Maitland - Ndabeni (down) and Southfield - Heathfield (down). The five corridors were selected as they were independently assessed to have varying states of condition, which is described later in the chapter. As condition information was sometimes only available for the up or down line of the chosen corridors, it was especially imperative that failure location was identified for each line. Additional repairs for failure modes that do not form part of the events log were added to the failure history of each corridor from the work orders database. No extensive renewals were reported for any of the corridors. Each perway corridor trend analysis was conducted from the first failure in the time envelope that data was recorded in. The time intervals for each perway corridor are presented in Fig. 5.1. First failures for Philippi and Maitland occurred fairly late in the recorded time envelope. The time from the previous failures was uncertain as there was some doubt that a failure could not have occurred for such a long period of time for each instance. The inter-arrival time for such a failure in the Maitland case would be 1132+ days, which is almost twice any

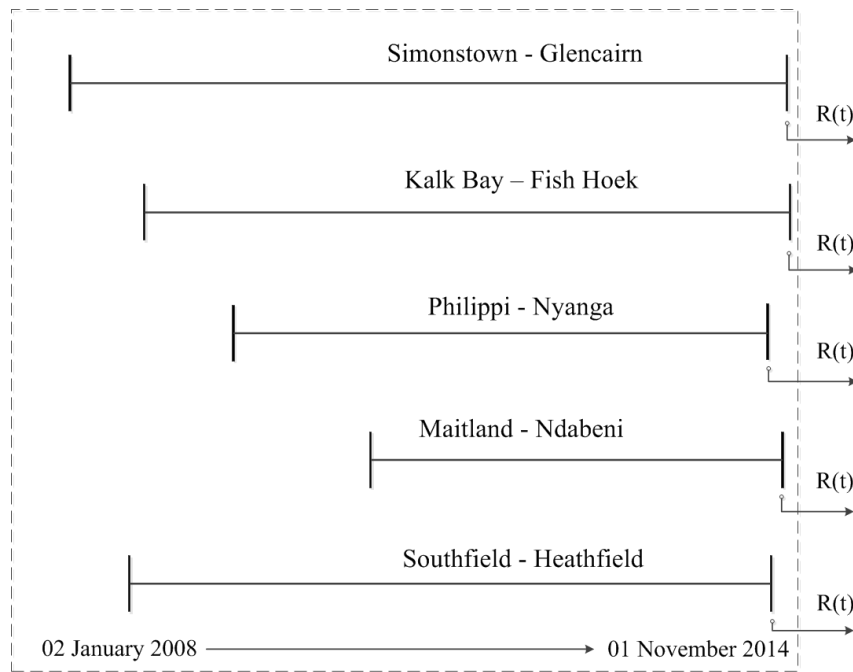


Figure 5.1: Time intervals over which trends were established for five perway track corridors.

other inter-arrival time recorded for that system. In addition, analysing the failure trend without being able to assume a failure at the beginning of the period skews trend tests as the first arrival time cannot be accurately specified, especially for such a long initial arrival time. For these reasons, the system was only analysed from the first recorded failure for these two perway corridors. The figure depicts times at which reliability prediction began for each system. To ensure consistency between each corridor for the purposes of analysis, it was decided that the time period for reliability predictions is relative to the time of the last failure in the perway corridors. In other words, systems are not compared in real time, but rather by their reliability performance after their last recorded failure. This time is denoted $+t$. The end times for system analysis were all within one month of each other thus old reliability trends were not compared with new trends.

Each of the five perway corridors is a repairable system. Fig. 2.16 is consulted to guide the evaluation of model determination for these repairable systems. According to the figure, data first needs to be tested for trend to estimate its dependence and determine whether it displays identical behaviour or not. The Laplace trend test was conducted, which only provided assurance that one of the five systems displayed a trend (section 2.11.1). A second test, the Lewis-Robinson test was conducted and from this test, a trend was identified for the other four perway corridor failures. Other trend tests such as the

Mann and Military Handbook tests were not needed for this analysis. All five systems displayed an increasing trend. Table 5.1 summarises these findings. High U_{LR} statistics from the table are evident due to the large variance experienced by each dataset. The large variance confirms that data is not identical and will therefore tend away from a HPP model. The least squares for each system produced the parameters discussed in the table and revealed a preferred form of the NHPP model, according to the coefficient of determination. Both power law and log linear models were made useful in this analysis.

Table 5.1: Trend testing and curve fitting for five perway corridors.

Perway corridor	U	U_{LR}	Reliability characteristic	Model	Parameters
Simonstown	+3.13	-	Degradation	Log-linear	$\alpha_0 = -8.269$ $\alpha_1 = 1.9 \times 10^{-3}$
Kalk Bay	+1.11	+3.34	Degradation	Log-linear	$\alpha_0 = -6.739$ $\alpha_1 = 1.1 \times 10^{-3}$
Philippi	+1.51	+4.44	Degradation	Power law	$\lambda = 1.1 \times 10^{-5}$ $\beta = 1.829$
Maitland	+1.11	+4.37	Degradation	Power law	$\lambda = 7.2 \times 10^{-6}$ $\beta = 1.874$
Southfield	+1.32	+3.84	Degradation	Log-linear	$\alpha_0 = -7.830$ $\alpha_1 = 1.3 \times 10^{-3}$

Reliability against time plots for each perway corridor are compared in Fig. 5.2. Remember that each system is in its own time window but these systems are compared on the same time axis for the purpose of maintenance prioritisation. The graphic illustrates a spectrum of reliability degradation from the poor Simonstown case to the seemingly more stable Southfield corridor. The difference in reliability for perway corridors is an indication that there is room for maintenance prioritisation to improve corridors with weak reliability performance. This spectrum is contained within reasonable limits which is an indication that the analysis has strength. It is important to understand how long it takes each system to degrade to 90% reliability as operations requires perway to perform with 90+% reliability. Whether this expectation is realistic or not considering present perway condition is a separate matter. The Simonstown - Glencairn corridor degrades to 90% reliability in 3.3 days. This is followed by 7.4 days for the Kalk Bay - Fish Hoek corridor, 10.2 days for Philippi - Langa, 10.5 days for Maitland - Nyanaga and 16.2 days for Southfield - Heathfield.

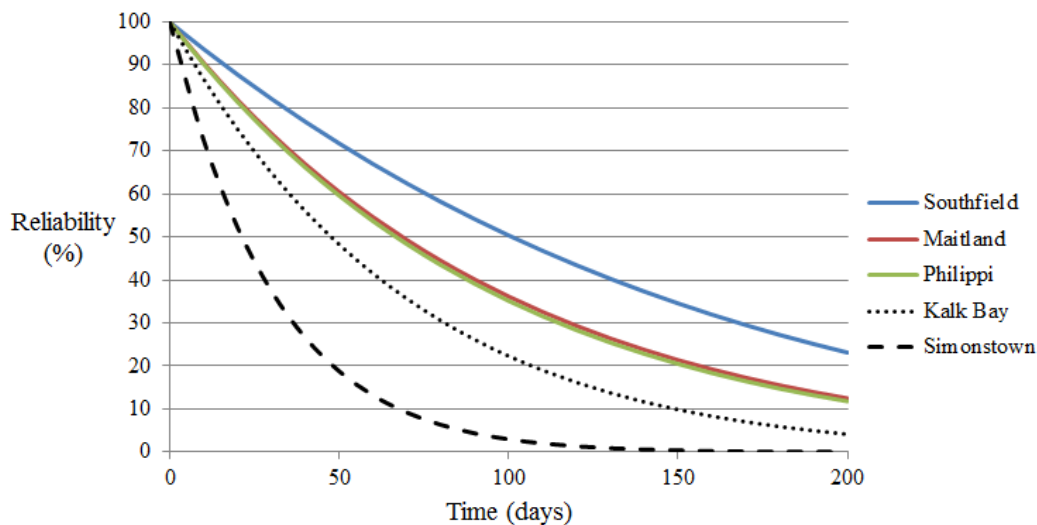


Figure 5.2: Reliability against time for five perway corridors in the Metrorail Cape Town network.

5.2 Risk analysis of multiple perway corridors

The risk metric for each of the five perway corridors is calculated in the same way as the Simonstown - Glencairn corridor from section 4.4. The Simonstown corridor metric was calculated as $77.8 \text{ mins}/3.3 \text{ days} = 23.6 \text{ mins/day}$. Following, Kalk Bay at $103.5 \text{ mins}/7.4 \text{ days} = 14.0 \text{ mins/day}$; Philippi at $79.7 \text{ mins}/10.2 \text{ days} = 7.8 \text{ mins/day}$. Maitland is $70 \text{ mins}/10.5 \text{ days} = 6.7 \text{ mins/day}$ and finally Southfield at $75 \text{ mins}/16.2 \text{ days} = 4.6 \text{ mins/day}$. The failure count for each corridor is dominated by failures that cause less delay, as the asset failures that caused these delays generally have a shorter lifetime. For instance, a rail life is approximately twenty years but a block joint's life is much shorter, causing more frequent service disruptions due to an asset like broken block joints. In this way, there is enough variation in the average delay of each system to observe that some of the systems have long delay assets that fail frequently. These systems are surely severe cases for maintenance that should be addressed more urgently. One must be careful not to assume that because the average delay for the corridor is high, that the corridor experiences many high severity failures. For instance, a corridor may have experienced four rail breaks and one section with slack. This might indicate that the section contains rails that were installed with poor quality but the rest of the corridor may operate efficiently. The combined effect of reliability and average delay is presented in Fig. 5.3, in which the five corridors are ranked in the risk matrix.

The calculated risk metrics fell into three out of the four risk categories. The last, unused risk category (green) was designed for systems that exhibit reliability improvement or that have not experienced enough failures to express

1/2 +	10	20	30	40	50	60	70	80	90	100
1/4 +	5	10	15	20 X	25	30	35	40	45	50
1/6 +	3.33	6.67	10	13.3	16.7	20	23.3	26.7	30	33.3
1/8 +	2.5	5	7.5 X	10	12.5 X	15	17.5	20	22.5	25
1/10 +	2	4	6	8 X	10	12	14	16	18	20
1/12 +	1.67	3.33	5	6.67	8.33	10	11.7	13.3	15	16.7
1/14 +	1.43	2.86	4.29	5.71	7.14	8.57	10	11.4	12.9	14.3
1/16 +	1.25	2.5	3.75	5 X	6.25	7.5	8.75	10	11.3	12.5
1/18 +	1.11	2.22	3.33	4.44	5.56	6.67	7.78	8.89	10	11.1
1/20 +	1	2	3	4	5	6	7	8	9	10
	0+	20+	40+	60+	80+	100+	120+	140+	160+	180+

Expected delay per incident for a track corridor (mins)

16+	Tamping, replacements, PM, <u>inspection across entire corridor, planned maintenance.</u>
8.5+	Tamping, replacements of critical failure modes, PM of critical failure modes, planned maintenance.
4+	General inspection, PM and replacement of faults found.
0+	General inspection, do maintenance as previously.

Figure 5.3: Risk matrix for the comparison of five perway corridors in the Metrorail Cape Town network.

a reliability trend. One may argue that the more years, for which data is available, the less systems will fall into the green category. In reality, systems are expected to be refurbished on a five to ten year time horizon which limits the number of failures available for trend analysis. A condition estimate for the track corridor must correlate with systems that exhibit less than four failures, the minimum needed for trend analysis.

As expected, the Simonstown corridor presents the highest risk. This risk category identifies perway corridors for which train service should not continue. After observing the failure modes for the corridor; derailments, broken rails, slacks and condition problems were encountered. This means that the poor reliability experienced cannot simply be solved by maintaining assets to prevent one or two critical failure modes. The only way to replenish the survivability of this system is to freshly install new perway. After the analysis was completed, the researcher discovered that a maintenance decision was made by Metrorail Cape Town to close down the Simonstown - Glencairn perway and refurbish it, as predicted by the risk matrix. This replenishment began in 2015 after 1 November 2014, the time when the failure observation period ended. This event is a real life example of the use of the risk matrix. The Kalk Bay - Fish Hoek corridor fell into the yellow risk category, which is focused on mitigating critical failure modes to reduce system risk. Here, components relating to the failure mode are fully serviced. The failures experienced were mainly caused by gauge, kickout and block joint problems. As gauge and kickout are problems related to rail fasteners, the potential maintenance solutions were narrowed down. Block joints are fairly standard to service. These types of failures justified the specification of the risk category and explains the effect of reliability and severity on levels of risk in a practical maintenance environment.

When it was observed that Maitland - Ndabeni and Philippi - Nyanga corridors obtained similar risk metrics, the same failure characteristics were expected from each system. Both corridors experienced mixed failure modes, with dominant failures occurring due to broken block joints in each case. As Maitland and Philippi corridors are close to the dividing line between risk categories on the matrix, the question arises: ‘what should be done for borderline cases?’ As the risk matrix is a decision support tool, it is imperative that decisions made from the risk matrix are confirmed by evidence from observations. For example, it would be wise to conduct preventative maintenance on the block joints for Maitland and Simonstown, which is an action specified by the yellow risk category. On the other hand, if the maintenance schedule is too tight, Maitland and Philippi would be the first corridors to have preventative maintenance overlooked. As for Southfield, the sparse risk renewal category is encountered. When failures or faults are encountered during planned maintenance in this category, renewals are conducted. This is due to the low regularity of maintenance for these corridors. It is better to fully

renew damaged components when maintenance on the component is not likely to occur again soon. Failures for Southfield were centred around points at railway switches that needed to be packed to align the track correctly as well and broken block joints. This implies that new points blades might need to be installed at local failure sites. The majority of maintenance conducted in the green risk category is servicing, such as oiling parts and tightening loose connections.

The risk matrix does not only exist for preventative maintenance decisions but for condition based maintenance as well. The risk metrics allow sequential ordering of risk which provides a queuing solution for condition based maintenance jobs. The tamping of railway tracks is conducted on a priority basis. The current prioritisation tool for this kind of maintenance is TQI. The risk metric will then correlate with TQI to gain additional confidence for prioritisation of tamping jobs. The risk metric changes the definition of where tamping is necessary by suggesting that tamping may be required to reduce risk, even if the current track condition is not in a critical state. For example, the ballast in a perway corridor may be poor in localised areas, which causes slack of the rail between sleepers. This slack generates risk for the perway corridor, but the TQI of the corridor is fairly low relative to other corridors. In this situation, the risk tool picks up faults where TQI was not able to and it provides a clearer understanding of where maintenance is needed.

5.3 Validation of risk matrix

A validation has been performed in part in the previous chapter, by explaining how the matrix sheds light on current maintenance decisions for five perway corridors. It was further shown that recommendations from the risk matrix were accurate for the Simonstown - Glencairn corridor, as a full renewal plan is in action. For further validation, it is shown that the risk matrix has a relationship with the TQI metric that is used to assess the condition of perway corridors. When the five systems were analysed, failure modes caused by anomalies that are external to perway degradation were excluded from failure analysis. For cases such as rail clip theft and sand on tracks, it is possible to geographically plot these failure modes to determine which areas experience the most service delay as a result of these factors. The locational model described is a separate model from maintenance management of perway deterioration. As a result of the exclusion of external failure modes, the condition of a given corridor should more closely correlate with the risk of the corridor. This is because perway condition degradation occurs in conjunction with an increase in the number of failures of perway components. For this reason the five chosen corridors were analysed as they are extremes cases in the TQI spectrum. The TQI values for each corridor are presented in Table 5.2, remembering that 7.5

is the highest TQI value recommended for mainline tracks (Zaayman, 2013). TQI statistics were measured bi-annually between March 2009 and September 2013. This is not the exact time envelope over which failure data was captured but it does represent five of the seven years for which failure data was captured. TQI data from 2008 and 2014 was omitted due to lack of availability and an assumption was made that the TQI data is representative of the failure data as the sample size for TQI is not much different from that of failure data. In addition, the sample evenly spans across the population time envelope.

Table 5.2: TQI values and risk metrics for five perway corridors in the Metrorail Cape Town network.

Corridor	Risk metric (mins/day)	TQI low	TQI average	TQI high
Glencairn - Simonstown	23.6	17.9	19.5	22.2
Kalk Bay - Fish Hoek	14.0	10	10.4	11.1
Philippi - Nyanga	7.8	7.6	8.3	9.1
Maitland - Ndabeni	6.7	7.3	8.6	9.4
Southfield - Heathfield	4.6	6.1	6.4	6.7

Table 5.2 reveals high TQI values and low TQI, which are the highest and lowest recorded TQI values for the corridor in the five year period. The risk metric values decrease proportionally with TQI. The relationship is not perfect between these metrics as there are obvious differences between them but it does confirm that risk metrics are in proportion to expected values. Southfield experiences some of the best TQI measurements in the Metrorail network but these values for TQI are rather high by the industry standard. The TQI values have increased as the system has aged, which confirms that the network is in a critical maintenance state. The risk matrix is designed for maintenance prioritisation for this reason, to eliminate critical risks to shift away from fire-fighting type maintenance. TQI not only provides a benchmark for risk metrics, but additionally has sibilance to maintenance decisions made from risk metrics. As the risk matrix is a decision support tool, it was not designed to replace TQI but rather come alongside it to inform decisions made and create fresh angles from which problems could be viewed. If it was the same as TQI, it would be obsolete and if it didn't correlate in some way, it would be invalid. Decisions from both tools overlapped for the Kalk Bay - Fish Hoek corridor. The gauge TQI parameter reached a value of 2.0. There are five geometric parameters for TQI and gauge is expected to be three times less than other parameters as it represents heavily constrained geometry (rail clips, sleepers and ballast). Rail clip fatigue is a typical cause for gauge deviation

as the track is not being fastened with the required force. This is confirmed by Sadeghi and Askarinejad (2009). In this case, the risk-based decision to maintain the critical gauge and kickout failure modes and thus replace rail clips is the same decision that results from TQI measurement. The risk matrix method is, in this way, validated both numerically and by case.

5.4 Sensitivity analysis of selected risk models

Although NHPP model selection is supported by statistical practise and literature, it is helpful to understand what reliability outputs are generated if only a log linear or power law model was selected for all perway corridors under analysis. This exercise depicts how sensitive reliability outputs are to best model selection. Table 5.3 presents reliability and risk metrics for log linear and power law models. These metrics are compared with the general TQI trend between perway corridors to see the effect of model selection on validation. When comparing reliability metrics for both models, Simonstown and Maitland have large variation in comparison to other perway corridors. The log linear model always degrades faster than the power law model for each corridor, which can be understood from a ROCOF graph. The log linear model accurately captures the extremes of risk for different perway corridors but risk for Maitland is far higher than expected. In proportion to Kalk Bay, Maitland should have a lower risk metric if perway condition is indeed sensitive to risk variation. The power law model for Maitland provides a better fit to observed data and a large error exists when the log linear model is chosen in this case. The power law model has a better correlation with TQI but does not capture extremes for risk well. The risk value for Simonstown should be higher, considering that it is in a state of disrepair. The sensitivity analysis shows that model selection is important to determine accurate results.

Table 5.3: Sensitivity for risk and reliability comparison of log linear and power law models.

Perway corridor	$T@R_{90\%}$ log linear (days)	$T@R_{90\%}$ power law (days)	Risk log linear (mins/day)	Risk power law (mins/day)	TQI average
Simonstown	+3.3	+6.7	23.6	11.6	19.5
Kalk Bay	+7.4	+9.0	14.0	11.5	10.4
Philippi	+7.0	+10.2	11.4	7.8	8.3
Maitland	+4.2	+10.5	16.7	6.7	8.6
Southfield	+16.2	+20.1	4.6	3.7	6.4

Chapter 6

Conclusions and Recommendations

In this chapter, a summary of the research findings, including successes and limitations is presented. The original research problem is discussed as well as solutions which met the set objectives for the project. Conclusions enabling success over the research limitations are remembered as these roadblocks will be experienced by others in the field. The contribution of the research to theory is mentioned so that others may not repeat advancements made. Also, recommendations are provided to help determine who may benefit from the research outcomes and how they may use these outcomes. Finally, forward movement on the topic is outlined and new topics that could help to improve deteriorating infrastructure in aged railway environments are presented.

6.1 Back to the research problem

Are there any maintenance tools that quantify the risk and reliability of a section of perway for the sake of maintenance prioritisation in the South African passenger railway industry? The stated research problem was answered through the creation of a reliability-based risk matrix decision support tool, which grades perway corridors between train stations according to the calculated risk that the corridor presents towards service delay. This tool is effective for maintenance prioritisation as it grades corridors into four risk categories as well as providing an ordered list which describes which corridors are in urgent need of maintenance, compared to others. In fulfilment of the strategy, the following objectives were met, so as to overcome the research problem:

- *To construct a reliability model representing the probability of successful operation of the train service from the perspective of a section of perway and populate the model using quantitative statistical failure data. A reliability model for perway was selected after deciding between component*

level or asset level for analysis. Failure data was extracted for perway corridors and trends were sought for data, consulting repairable system reliability. Point process models were selected for corridors, supported by goodness of fit methods. Reliability was calculated for each system from the end of recorded failure history using the model platforms provided.

- *To create a reliability-based risk model that compares the risk of severe service-stopping failures of different perway sections for the purpose of perway section maintenance prioritisation.* Failure modes for perway were identified from a FMECA. Operations reports on each failure mode were consulted to obtain estimates for average network delay. An expert in the field of railway perway was consulted to adjust delay estimates for accuracy. Failure data was used to generate reliability for perway corridors as the likelihood component of a risk matrix. The severity component of the risk matrix is the average network delay from each perway corridor. Maintenance prioritisation is possible from this model due to risk categories as well as the order of criticality, according to risk metrics.
- *To validate the developed reliability-based risk model by comparing it to an appropriate condition-based tool currently used to make maintenance decisions.* Data for the condition-based tool was available for perway corridors. This TQI statistic provided a condition evaluation of each perway corridor. For each perway corridor, the calculated risk metric was compared to the TQI statistic and a correlation existed between the TQI and risk range for corridors. This correlation validates the relation between one risk measure and another, ensuring that the priority order of risk metrics is accurate. Maintenance decisions made from TQI values were similar to those made from risk metrics.

6.2 Theoretical contribution and practical application of methods

The researcher created new models and methodologies during the research process that are documented. These new models and methodologies are, notably, a reliability block diagram for reliability analysis of short sections of perway; literature on the comparison between TQI and perway reliability; a FMECA methodology as well as an FTA methodology. A summary has been provided for the step-by-step application of repairable system reliability methods to fit NHPP models to systems. No concise literature has been presented on this topic, although none can surpass the mastery of Ascher and Feingold (1984). A perway corridor reliability model was created, which takes into account train traffic decisions. A decision support risk matrix was created which

aids perway maintenance decisions, to be used in conjunction with TQI. These contributions are made available to help those involved with engineering asset management and statistical reliability analysis.

In practise, maintenance information pertaining to assets is logged into an asset management database called Empac within the PRASA environment. This software tags assets, keeps tabs on disposals and facilitates ordering of new parts. To plan maintenance based on this data, the data has to be extracted from the database in raw form and statistical analysis needs to be conducted by an individual. The developed risk matrix is helpful because it is an output for statistical data that otherwise might not have been used accurately to make decisions. It is not in itself a complete maintenance management tool but it does present a visual understanding of the state of the assets. Cost saving and exact reliability improvement is not yet quantifiable from the matrix and these depend on the decisions that personnel make when using the matrix.

The developed risk matrix is applicable to a condition-based maintenance environment where scheduling is ordered by queues, ranked according to priority. The risk metric for each perway corridor enables this queuing to take place. Preventative maintenance tasks for perway are prioritised by risk category from the risk matrix and the order of the risk metric is secondarily considered to determine priority. As perway maintenance decisions are complex, information provides grounds for better decisions to be made. The risk matrix and failure history for perway provides additional but concise information so that more informed maintenance decisions can be made.

6.3 Conclusions about limiting factors encountered

The relationship between reliability and maintenance has not been well quantified in literature. Maintenance obviously affects system reliability but it seems to be very difficult to separate this effect from the performance of a system. For now, we assume that a system is combined with its past maintenance decisions and any decision made over and above previous decisions can improve the reliability by a quantifiable amount. It is convenient to assume that sparse maintenance events have not affected system reliability much. As a pro-active strategy to improve maintenance, it is suggested that maintenance is conducted on a cluster of components so that the improvement of reliability can be quantified in the system. Suggested further research may enable management of reliability levels through specific maintenance tasks.

Reliability analysis is limited by the number of failures available for a system. Some long lasting components in other industries will, therefore, never have data available for the sake of analysis. These limiting factors were particularly prevalent in PRASA's maintenance environment as they limit the development of even better reliability models. The way to get around this problem is accurate data-recording and the availability of ample data for a system. For instance, if the condition of every sleeper was known for an entire perway network then failures and reliability could be quantified in an entirely new way. In conclusion, there is currently sufficient data available to populate the developed model, but with more data, new methods can be established to improve reliability predictions.

As the cost of train delay has many different interpretations including customer satisfaction, it is difficult to calculate. To reduce train delay is to improve asset up-time, which indicates that assets are being restored. The risk matrix does not necessarily reduce cost but encourages maintenance expenditure in an intelligent way, so that a financial backlog does not occur. This saves PRASA from an increasingly deteriorating asset fleet. The developed risk matrix is limited to theoretical application at this stage. For it to be used in a practical maintenance environment, the research methodology would have to be applied to many more assets on data from other sources to PRASA. If this was completed, the risk matrix would be able to instruct maintenance managers where it is most beneficial to allocate pre-defined resources. In addition, it could encourage increased expenditures when the current state of the perway network is critical in relation to desired reliability performance measures.

6.4 Future work

The research topic opens up a myriad of future research avenues to further the railway industry as well as statistical and asset management domains. The topic itself can be extended such that its application is in-line with future technology and problems can be solved in similar domains by applying the same methods.

6.4.1 Extending the research topic

Research has now been conducted on the quantification of reliability in the rolling stock (Conradie, 2015) and perway domains in the railway industry. Research needs to be extended to telecommunications and electrical domains within railway infrastructure so that reliability can be synergised through different interfaces of the railway service. Another extension of the presented research is the application of optimisation methods for the geographical place-

ment of maintenance crews. Operations research methods should be applied to improve route options for perway maintenance equipment to realise priority maintenance activities at the lowest cost.

6.4.2 New research avenues

The application of big data to the maintenance environments in the railway industry is a topic that the researcher feels has not yet been considered in literature. As data capturing becomes increasingly easier as a result of the decreasing cost of sensors and the enormous storage capacity and processing power of modern computers, it is important that data is essentially captured and effectively used. Big data removes the level of expertise that comes with standard data capture and makes so much information available that the manipulation of data is not required if one is to have a good picture of the observed environment. Training everyone to analyse data well in the railway environment would be impossible. Conversely, capturing data intelligently, with visual outputs, would give laymen tools to understand the state of their assets.

Appendices

Appendix A

FMEA Tables

A risk priority number is comprised of occurrence, severity and detection measures. Three tables are presented which are necessary to determine risk priority in an FMECA, namely Tables A.1, A.2 and A.3.

Table A.1: FMEA occurrence evaluation (Chin *et al.*, 2009), (Xu *et al.*, 2002).

Occurrence evaluation criteria	Likelihood of failure	Failure rate (average for 5 year period in Western Cape Region)	Ranking
Very high	Persistent failures	0.2 or more	10
		0.1	9
High	Frequent failures	0.05	8
		0.02	7
Moderate	Occasional failures	0.01	6
		0.005	5
Low	Relatively few failures	0.002	4
		0.001	3
Remote	Failure is unlikely	0.0005	2
		0.0002	1

Table A.2: FMEA severity evaluation (Chin *et al.*, 2009), (Xu *et al.*, 2002).

Severity evaluation criteria	Outage (Average delay minutes per incident)	Ranking
Extremely disrupting	More than 320 minutes	10
Very disrupting	More than 160 minutes	9
Very high	More than 80 minutes	8
High	More than 40 minutes	7
Moderate	More than 20 minutes	6
Low	More than 10 minutes	5
Very low	More than 5 minutes	4
Minor	More than 2 minutes	3
Very minor	More than 1 minutes	2
None	No discernible effect	1

Table A.3: FMEA detection evaluation (Chin *et al.*, 2009), (Xu *et al.*, 2002).

Detection	Description	Ranking
Not detectable	The risk is not detectable by existing control mechanisms in the system.	10
Almost undetectable	The risk is almost undetectable by existing control mechanisms in the system.	9
Very low	There is very low chance that the risk is detected by existing system.	8
Low	There is low chance that the risk is detected by existing system.	7
Moderately low	There is moderately low chance that the risk is detected by existing system.	6
Moderate	There is 50-50 chance that the risk is detected by existing system.	5
Moderately high	There is moderately high chance that the risk is detected by existing system.	4
High	There is high chance that the risk is detected by existing system.	3
Very high	There is a very high chance that the risk is detected by existing system.	2
Definitely detectable	The risk is definitely detectable by existing	1

Appendix B

Fault Trees for Top Ten Failure Modes

A fault tree analysis was conducted on perway failure modes. Fault trees for nine of the ten most critical fault trees are presented here, while the rail break fault tree is presented in the main body of the document. The rail clip failure mode is one of the ten critical failure modes but it is not presented as its root cause is captured in the rail break fault tree. The same situation exists for the skid mark failure mode. Fault trees are presented in B.1, B.2, B.3, B.4, B.5, B.6, B.7, B.8.

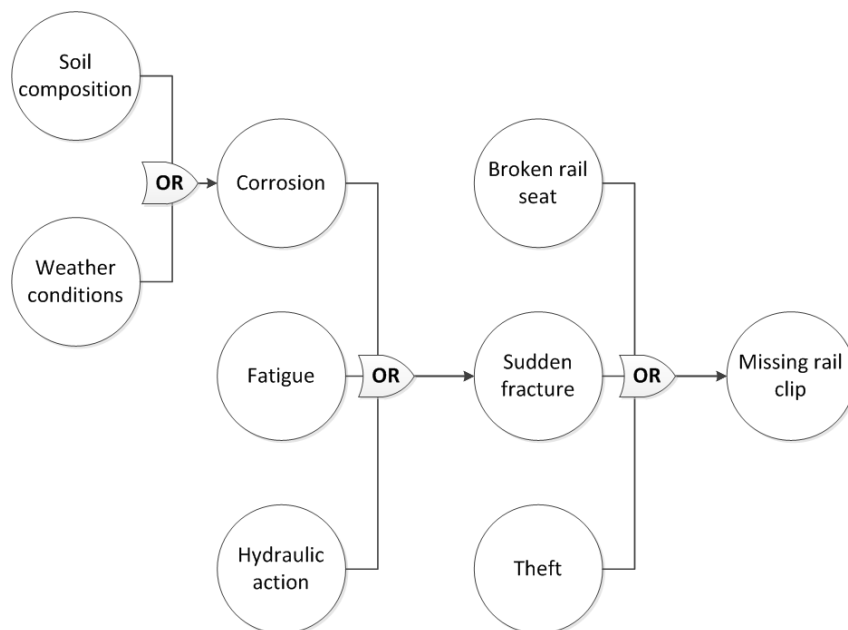


Figure B.1: Fault tree for rail clip failure mode.

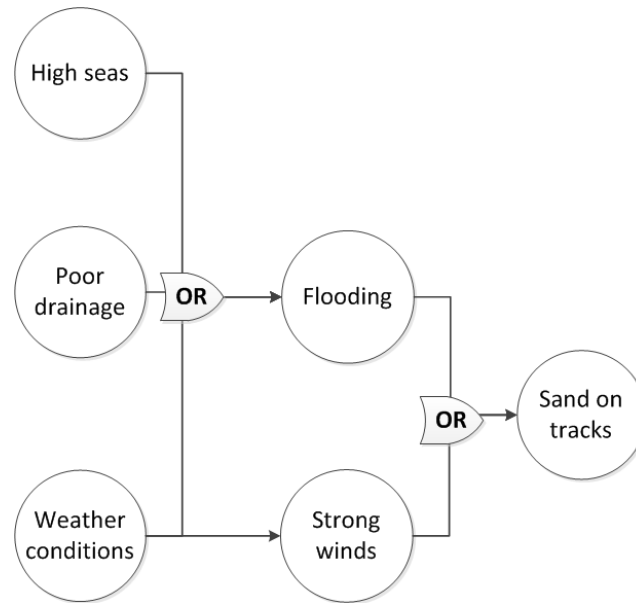


Figure B.2: Fault tree for sand on tracks failure mode.

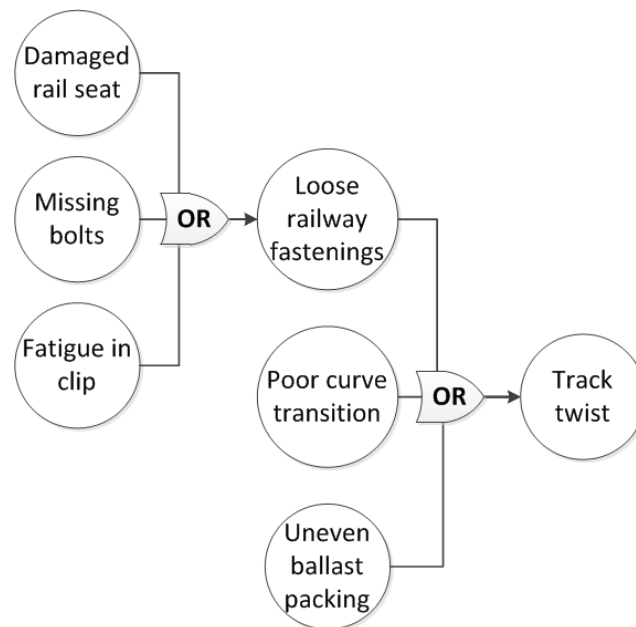


Figure B.3: Fault tree for track geometry twist failure mode.

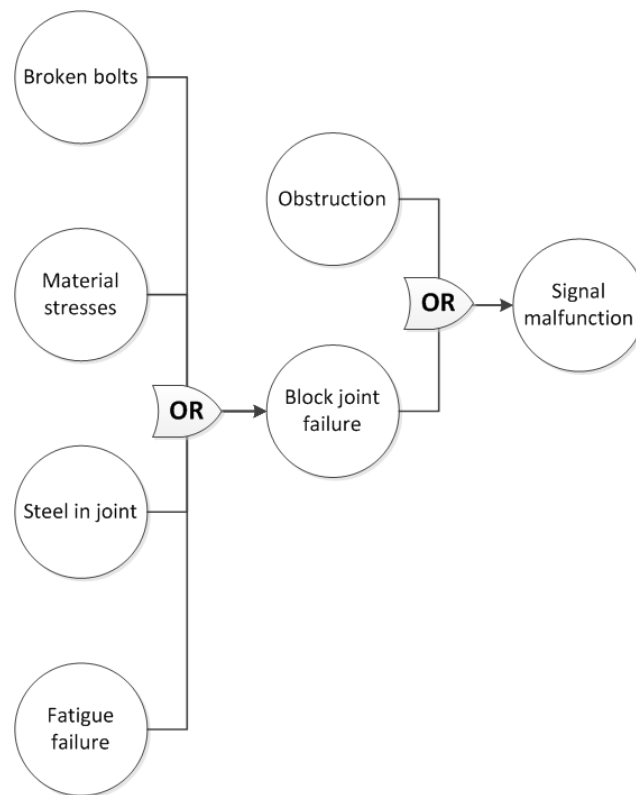


Figure B.4: Fault tree for block joint failure mode.

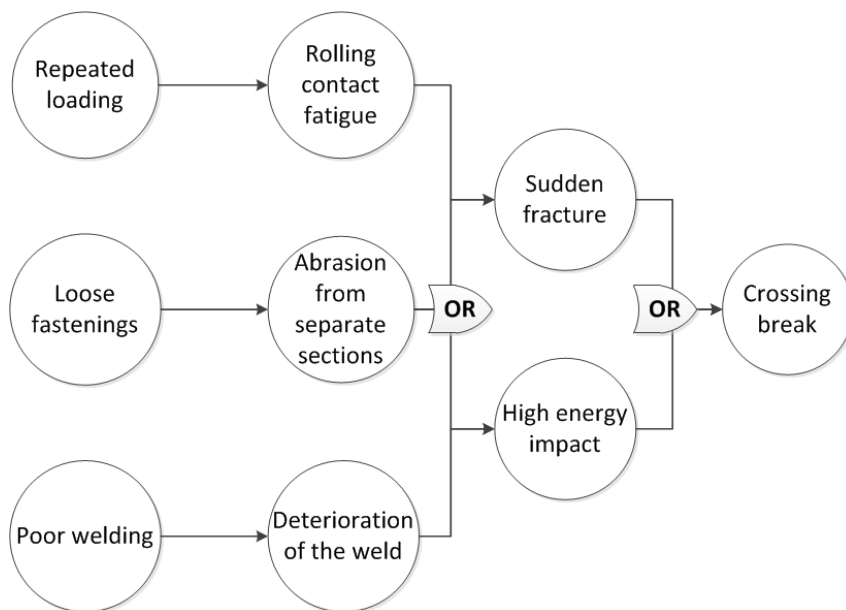


Figure B.5: Fault tree for railway crossing failure mode.

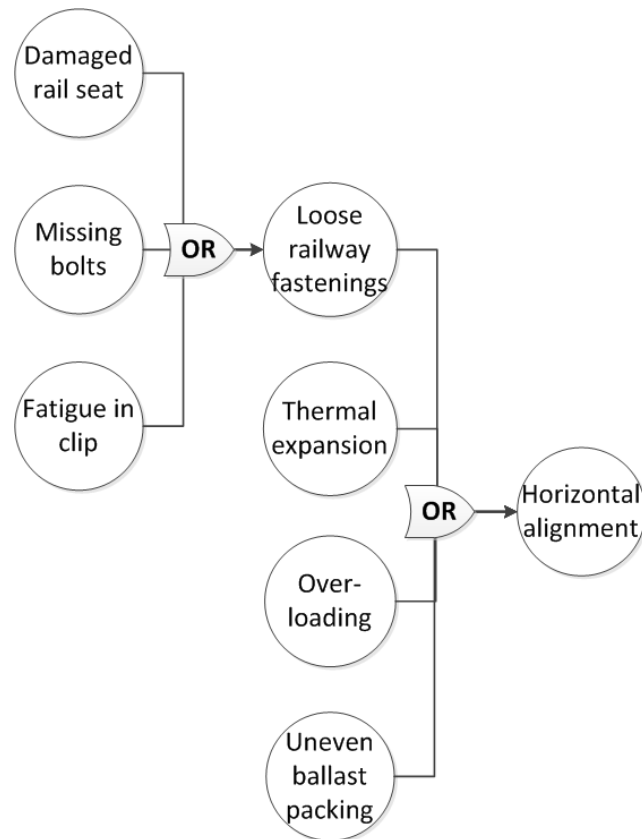


Figure B.6: Fault tree for track geometry horizontal alignment failure mode.

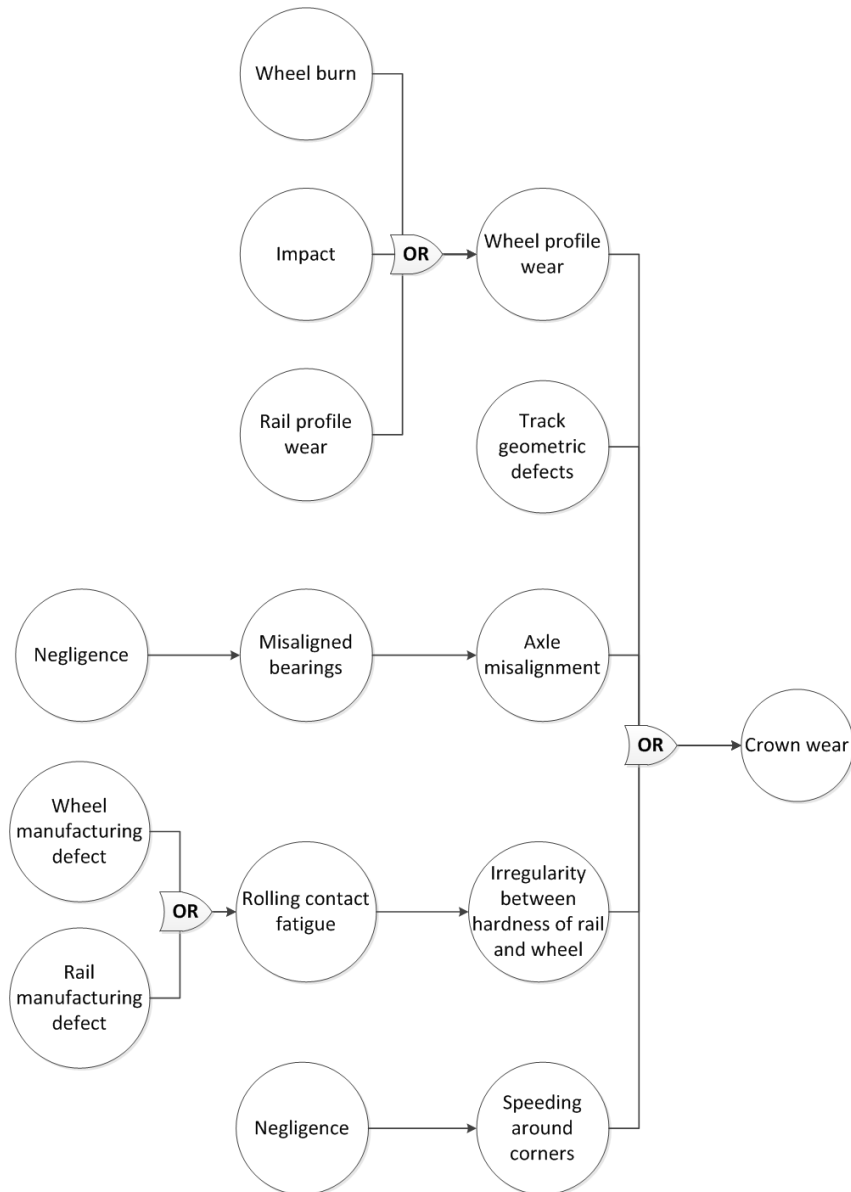


Figure B.7: Fault tree for rail crown failure mode.

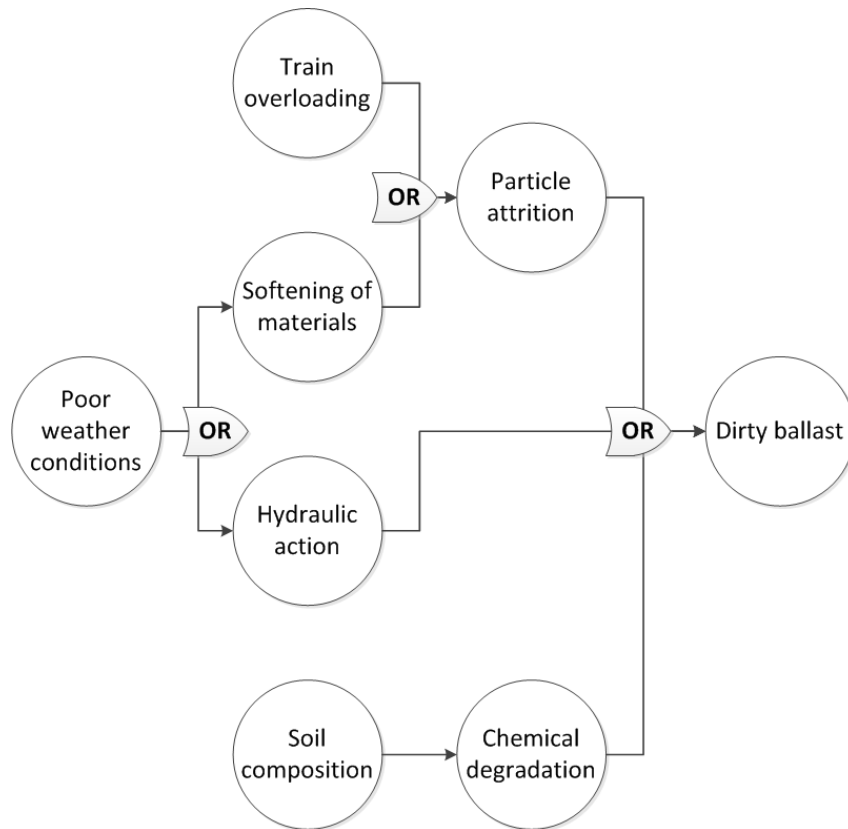


Figure B.8: Fault tree for dirty ballast failure mode.

Appendix C

Cumulative Failure Against Time for Five Perway Corridors

Cumulative failure against time graphs for four of the five perway corridors that were analysed in the main text are presented. Log linear and power law fits are included, as well as R^2 goodness of fit values. The fits are presented in Fig. C.1, C.2, C.3 and C.4. Table C.1 presents R^2 values for log linear and power law fits.

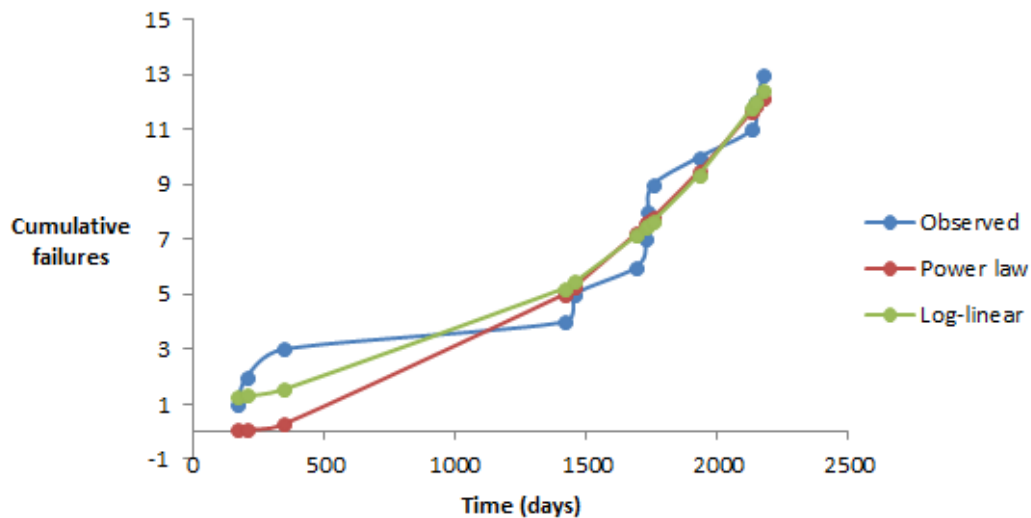


Figure C.1: Cumulative failure against time for the Kalk Bay - Fish Hoek perway corridor.

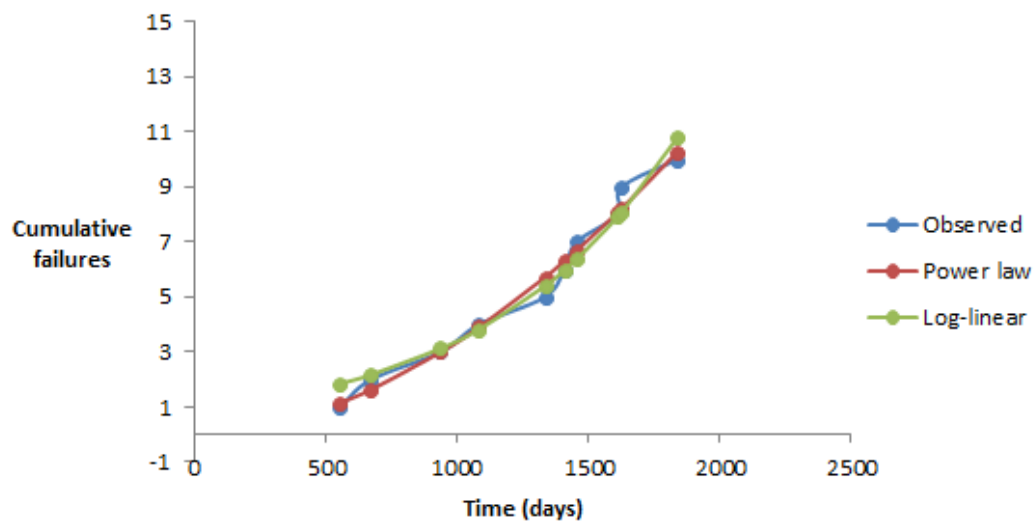


Figure C.2: Cumulative failure against time for the Philippi - Nyanga perway corridor.

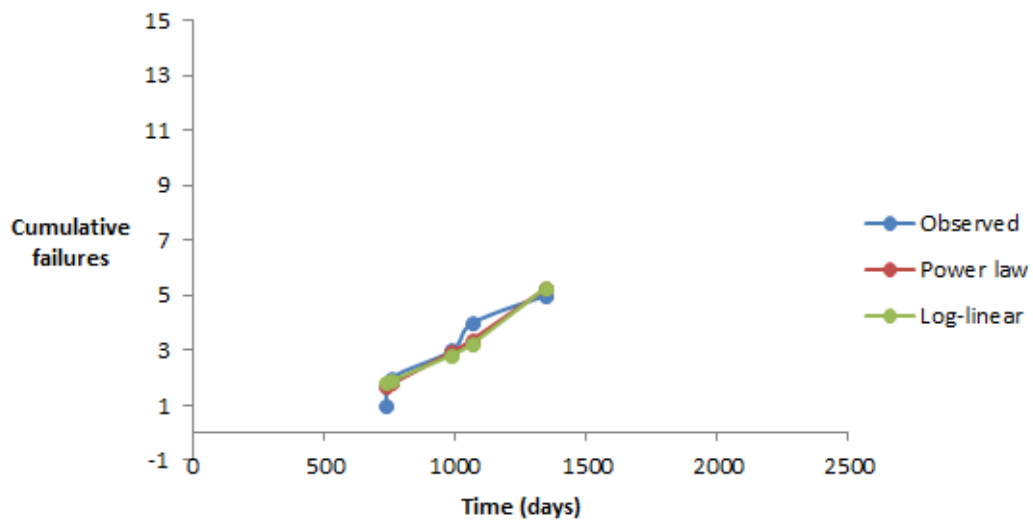


Figure C.3: Cumulative failure against time for the Maitland - Ndabeni perway corridor.

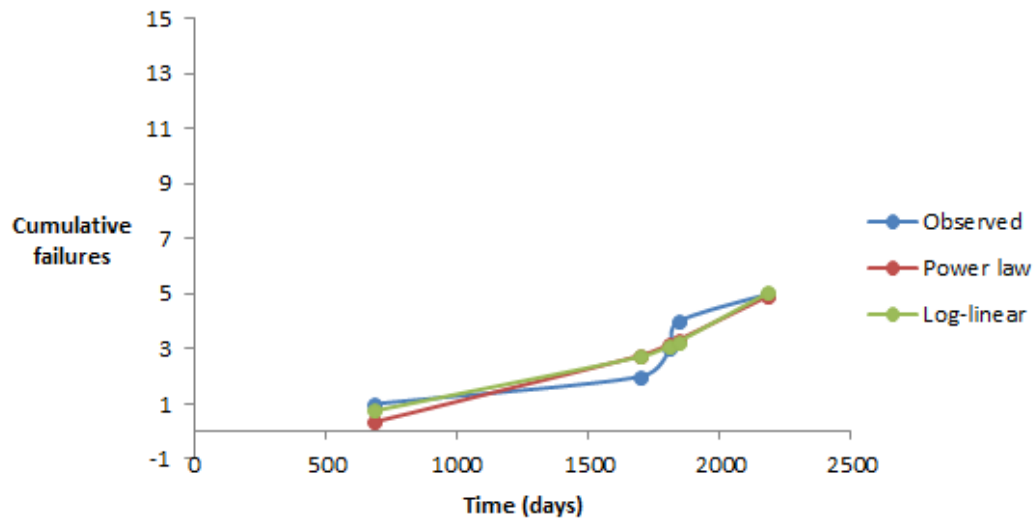


Figure C.4: Cumulative failure against time for the Southfield - Heathfield perway corridor.

Table C.1: Sensitivity for risk and reliability comparison of log linear and power law models.

Perway corridor	R^2 log linear (%)	R^2 power law (%)
Simonstown - Glencairn	94.2	89.8
Kalk Bay - Fish Hoek	97.2	94.5
Philippi - Nyanga	98.5	99.2
Maitland - Ndabeni	81.8	87.3
Southfield - Heathfield	97.1	96.1

Appendix D

Linear Estimation Method for Five Perway Corridors

A linear estimation method by Crowder *et al.* (1994) is applied to five perway corridors in support of R^2 goodness of fit values. These curves are presented in Fig. D.1, D.2, D.3 and D.4.

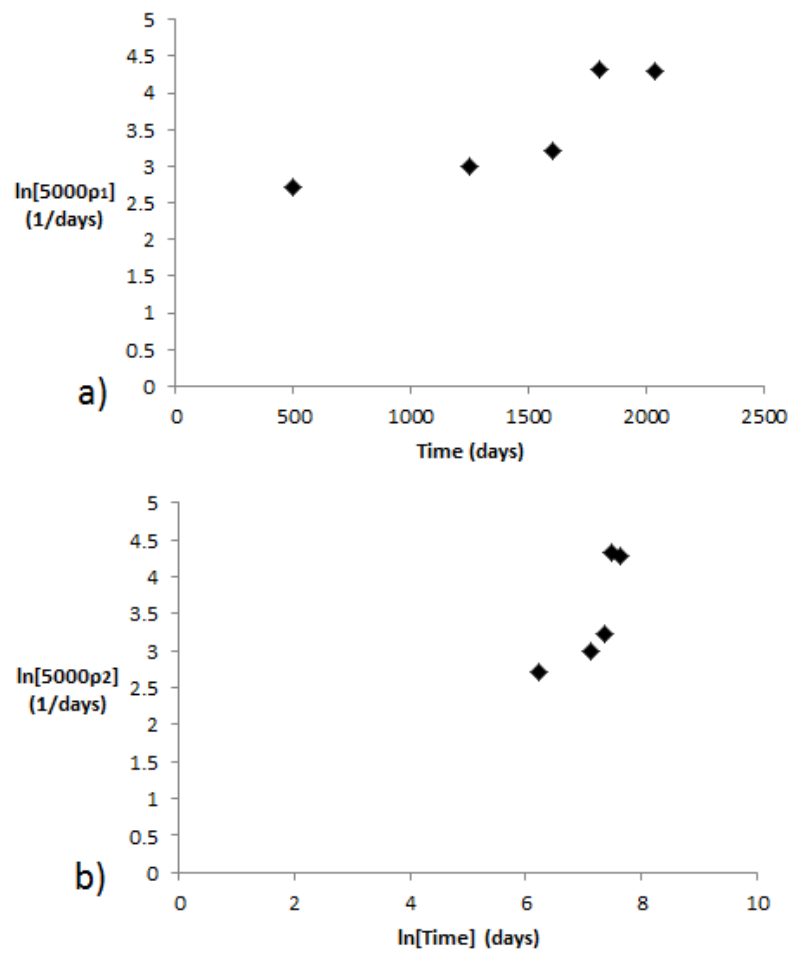


Figure D.1: Estimated ROCOF for Kalk Bay a) log linear and b) power law models

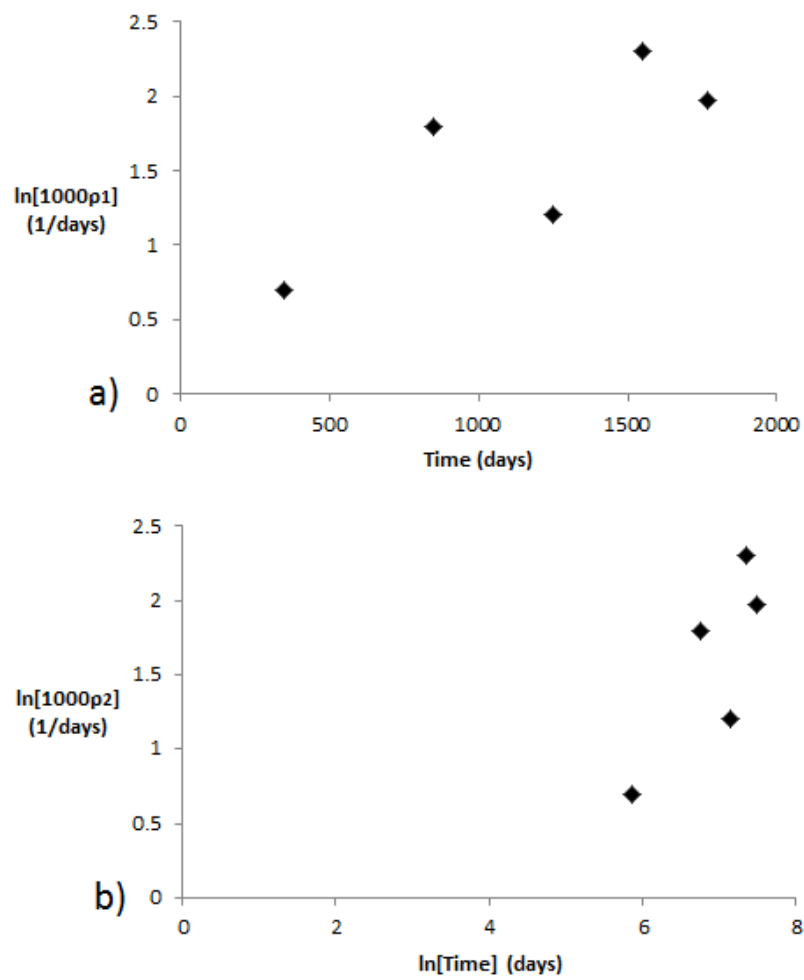


Figure D.2: Estimated ROCOF for Philippi a) log linear and b) power law models

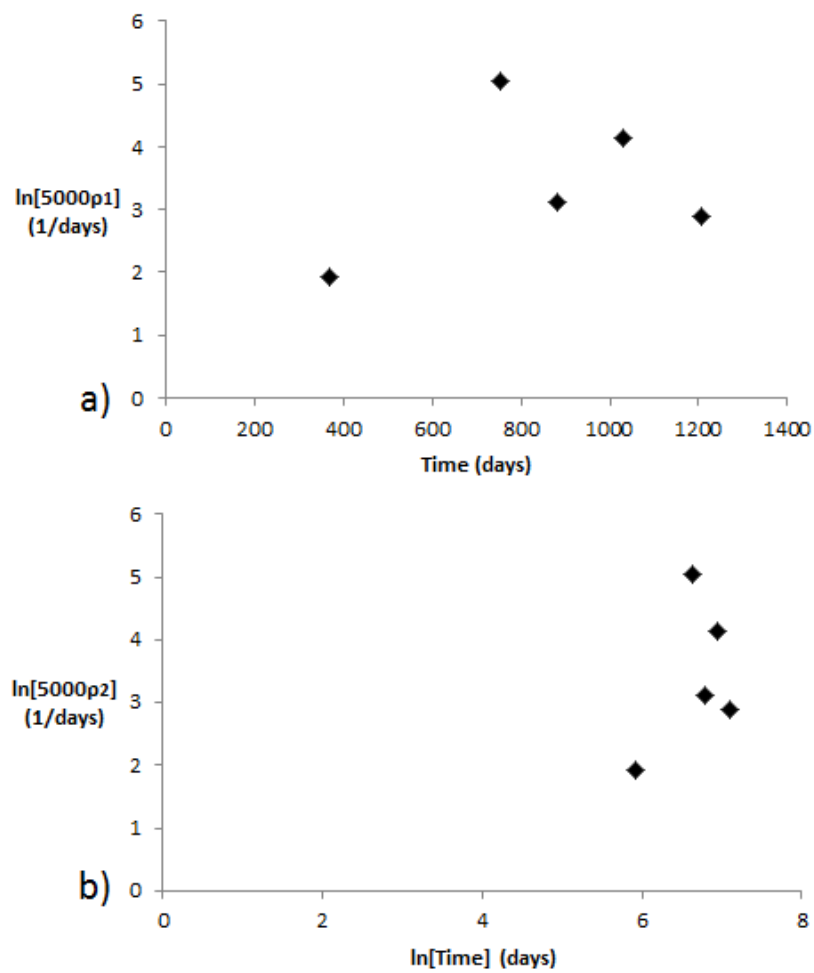


Figure D.3: Estimated ROCOF for Maitland a) log linear and b) power law models

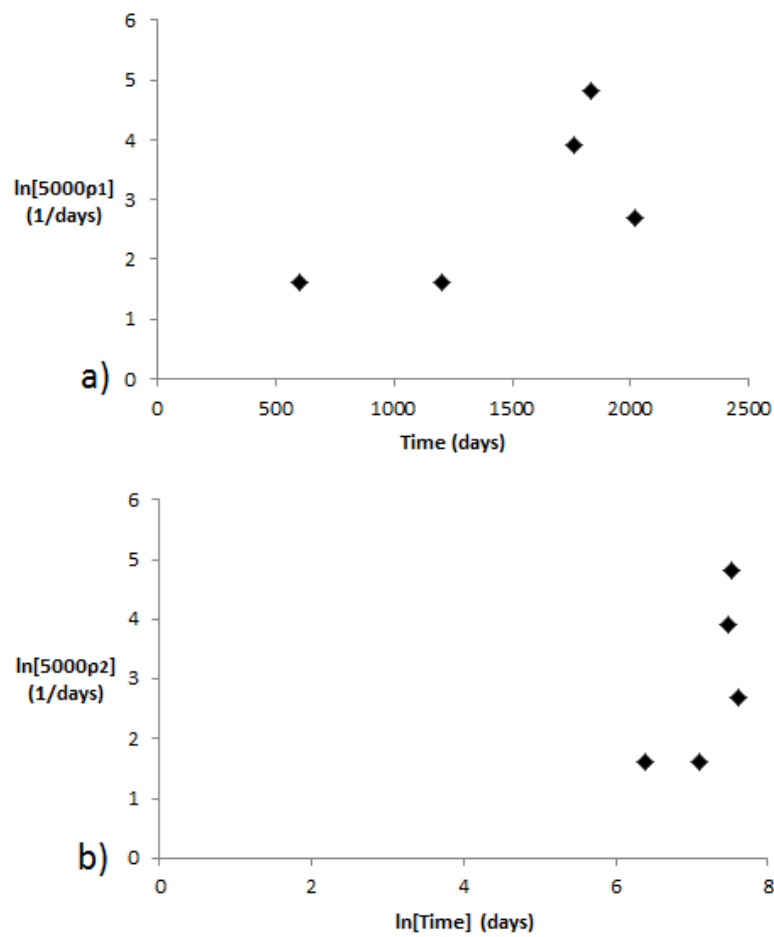


Figure D.4: Estimated ROCOF for Southfield a) log linear and b) power law models

Appendix E

Confidence Bounds on ROCOF for Five Perway Corridors

95% Confidence bounds are placed around the ROCOF for five perway corridors. Confidence bounds are plotted for log linear and power law models, namely Fig. E.1, E.2, E.3 and E.4.

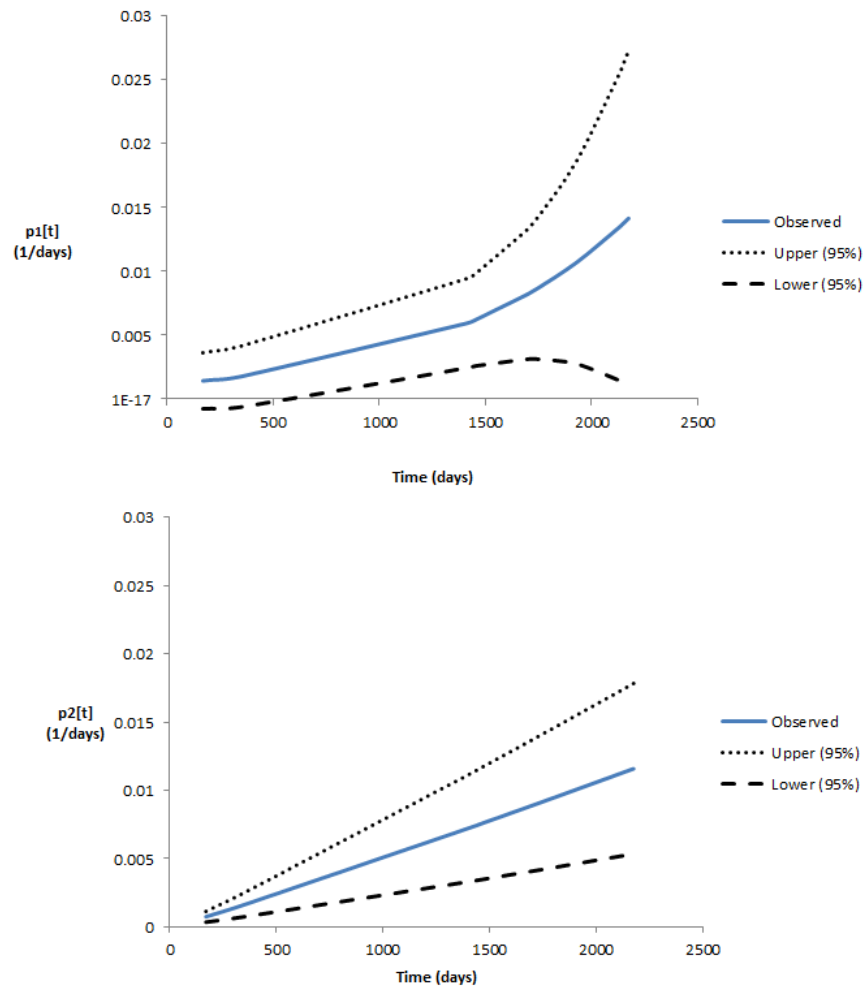


Figure E.1: ROCOF of the Kalk Bay perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models.

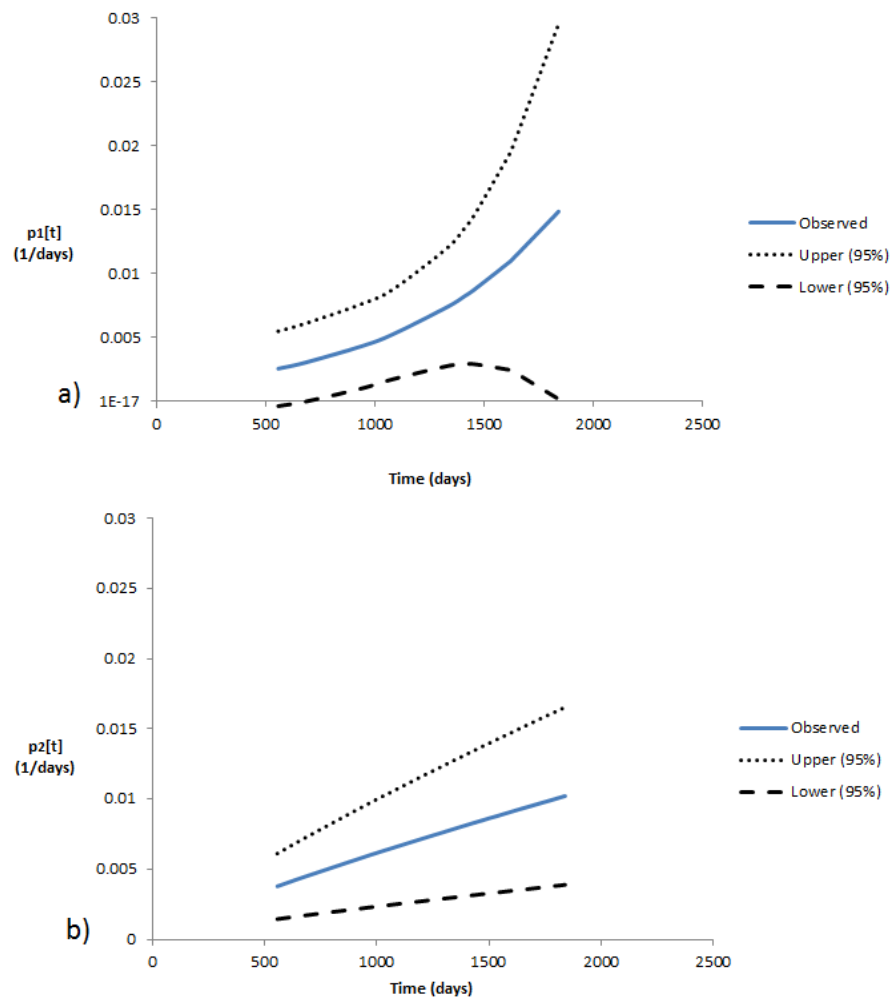


Figure E.2: ROCOF of the Philippi perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models.

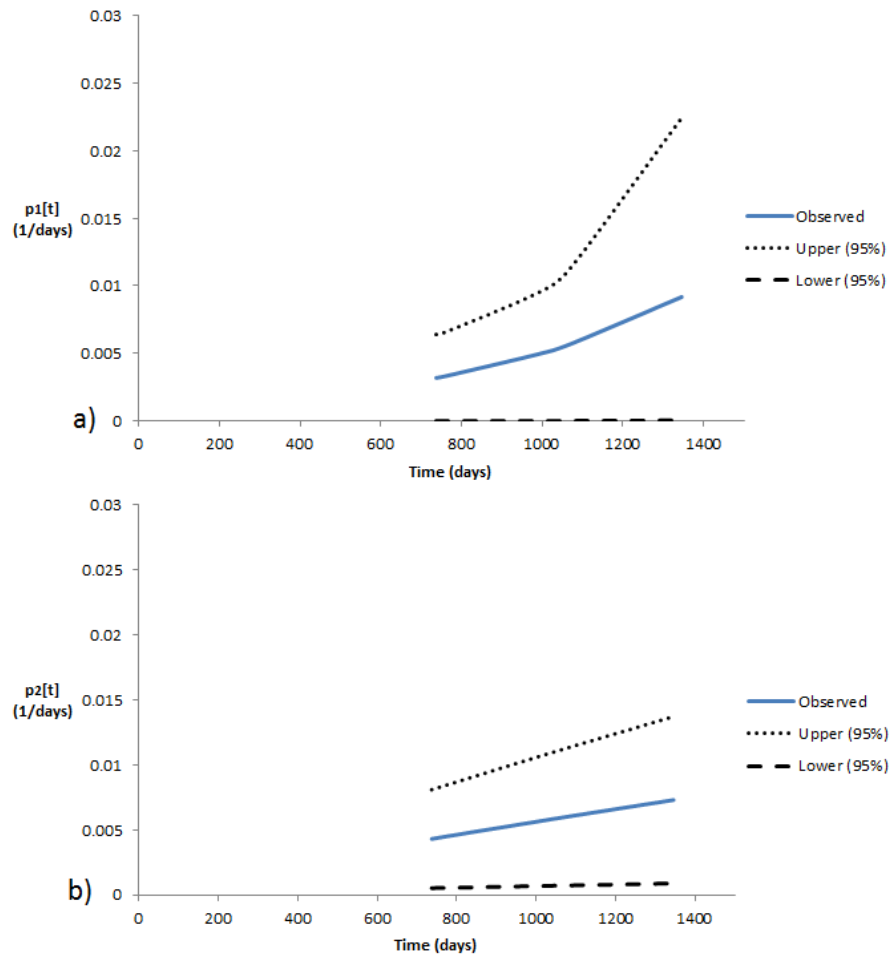


Figure E.3: ROCOF of the Maitland perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models.

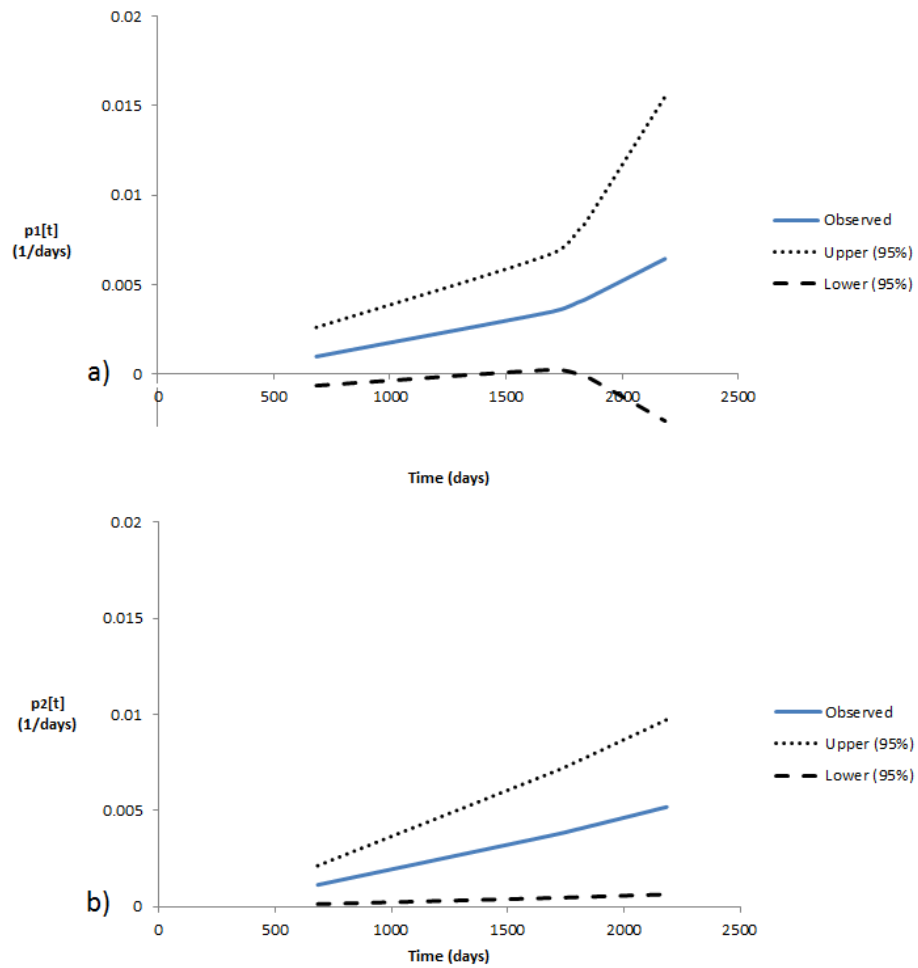


Figure E.4: ROCOF of the Southfield perway corridor with 95% upper and lower confidence bounds for a) log linear and b) power law models.

Appendix F

Urban Transport 2015 Conference Article

In June 2015, Mark attended the Urban Transport 2015 Conference in Valencia, Spain along with a colleague from the PRASA Engineering Research Chair. Mark presented a topic related to reliability in the railway perway context to an international audience. His paper has been published in the Urban Transport 2015 proceedings and the article can be found on Google Scholar.

A Review of critical problems from the desk of chief executive officers in the passenger railway service industry

M.D. Mc Naught¹, C.J. Fourie¹

¹*Department of Industrial Engineering, Stellenbosch University, South Africa*

Abstract

In the passenger railway service industry, there are a plethora of problems and every day, big decisions have to be made by executives as to which problems are most important. The key areas of a passenger railway service company are identified by conducting a customer needs analysis. These needs are then converted into operational requirements, which are considered to be the functional ‘building blocks’ of the company. An asset management cost analysis method is employed to determine three critical cost areas from the financial statements of ten passenger railway service companies. This gives executives a financial handle to engineering problems. The problems are discussed from an engineering perspective and a solution method is presented which will save cost in safety, maintenance and operations domains.

Keywords: railway, reliability, needs analysis, infrastructure, maintenance, safety.

1 Introduction

Passenger railway service companies, like other large companies, have multiple complex problems affecting the profitability of the company. These problems are the costs which hold the company back from excellent financial performance. These problems especially affect passenger railway service companies in developing countries who have not developed systems to effectively manage more than a few key problems, at any given time [1]. Thus, it is necessary to

develop a methodology which simplifies the multitude of problems experienced into a more manageable set of key performance affecting problems.

The developed methodology will allow effective communication between engineering managers and financial officers, which is where the handle is often lost on problems, due to prioritisation discrepancies between engineers and financial representatives [2]. The prioritisation of problems is also often lost in the complexity of trying to manage and run the business. Redirecting to the correct focus, the key identified problems are centred on customer needs. When these needs are satisfied, the business becomes more profitable. Thus, a tool is developed which identifies key problems which are not only costly to the business, but also important for the sake of customer satisfaction.

From a process of inductive reasoning and questioning, the following research objectives are set:

- Present finance-based evidence which justifies the severity ranking of different engineering problems.
- Determine a ranking of importance for compared customer needs.
- Conduct an analysis of problem topics from three critical problem domains.
- Present potential solutions for the sake of continued research.

2 Cost analysis method

A cost analysis method was designed to determine a critical improvement area in a developing passenger railway company (such as rolling stock quality control or infrastructure maintenance). The idea behind the method is to determine the basic functional ‘building blocks’ of the railway company, based on the needs of the customer. The building blocks or ‘costs’ are then compared on an operating cost basis, using the financial statements from ten railway companies.

2.1 Customer needs analysis

As the passenger railway service industry is service focused, the researcher argues that the ‘building blocks’ that make up the railway system should exist to fulfil customer needs.

Two interviews were designed; following the customer needs process developed by Takai & Ishii [3], to evaluate the customer needs of passengers in Railway Company A in South Africa. The procedure of the discussed customer needs development process is outlined. Sampled customers were first interviewed using a questionnaire to identify customer needs of the railway service. The customers were then asked to use the Subjective Clustering method (SC method) to group similar customer needs, which is discussed by Takai & Ishii [3]. The sample of customers was then asked to use the Affinity Diagram method, (AD method) which is the validation method for the grouping of customer needs [3].

The second interview was conducted, in which customers were given the refined needs and were asked to rank them based on importance. Customer needs were then ranked by the researcher, based on the customer collected information.

The interviews were conducted on an individual basis and account for customers who travel for both work and recreational purposes. The selected train station (Stellenbosch, South Africa) accurately captures a varying demographic as there are train users there who are students from the University, business men and woman from higher income and lower income classes and locals from the nearby township. An even number of customers from each demographic were interviewed. The interview results are presented in Fig. 1.

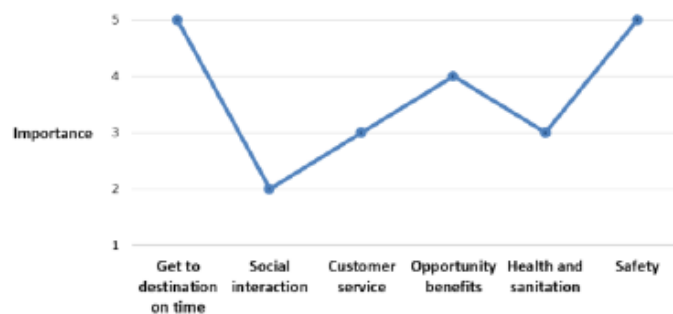


Figure 1: Importance diagram of refined customer needs

Fig.1 shows six clearly defined customer needs, which were cross correlated between the SC method and the dendrogram from the AD method [3]. The interviews revealed that safety and arriving on-time are the needs most important to the customer base.

2.2 Most basic system operational requirements ‘building blocks’

The needs from the customer base are converted into system operational requirements to determine key ‘building-blocks’ in the railway company. The format of the operational requirements is extracted from Blanchard & Fabrycky [4]. The operational requirements for a passenger railway service company comprised of: a mission definition, performance parameters, operational deployment, operational life cycle, utilisation requirements and environmental factors. The operational requirements reveal ‘building blocks’ which make-up the railway network from the business end. Fig. 2 illustrates a summary of the ‘building blocks’.



Figure 2: 'Building blocks' of a passenger railway network

2.3 Cost comparison from the financial statements

The 'building blocks' of the railway network are converted into 'costs' common to the financial statements (2012-13) of ten passenger railway companies, which are public documents. Any finance used to sustain a building block is considered to be a 'cost'. Each cost is compared on a yearly basis to ensure equivalency. For instance, capital debt is measured by interest expenses and loan repayments and land & buildings are measured by renewals, renovations and maintenance. Consistently high costs across different railway company's financial statements are considered to be critical costs, which is the focus point of this investigation. This method is similar to the life cycle costing method by Marquez et al. [5]. They compare costs on a lifetime basis, whereas this researcher compares costs on a yearly operational basis. Table 1 presents costs for ten passenger railway companies, including the totals for each cost and the totals for each company.

Table 1: Company 'costs' from 'building blocks' for ten passenger railway service companies

Cost (\$ millions)	Company										Total
	Amtrak	ARTC	China Railway Group Ltd	Deutsche Bahn	Ferrovie Italiane	Indian Railway	JR Central	Network Rail	PRASA	SNCF	
Capital Debt	567.00	28.16	14938.36	21092.04	20969.18	746.33	3914.58	48117.68	487.01	9447.38	120307.74
Customer service	101.67	-	121.01	888.50	1.74	0.03	110.98	140.99	24.35	846.23	2235.49
Infrastructure	772.67	185.18	20740.99	12264.11	1699.76	0.03	2059.98	2672.32	138.46	5110.80	45644.30
Land & Buildings	246.64	159.20	1798.77	5411.65	295.02	0.01	626.98	6340.03	45.27	2784.32	17707.90
Maintenance	205.20	-	-	27058.26	1475.11	-	-	951.00	78.84	-	29768.41
Operations	905.495	3.72	-	21493.3	-	-	7459.97	6308.33	264.58	1575.42	38010.82
Rolling Stock	877.57	-	-	6671.35	596.32	0.03	1653.48	-	232.31	-	10031.06
Safety	321.37	22.27	-	370.40	304.80	-	304.80	-	85.72	986.41	2395.74
Telecommunications	378.43	12.33	242.02	2739.01	3.49	0.18	1972.89	-	74.30	1126.59	6549.22
	4376.04	410.84	37841.15	97988.61	25345.41	746.61	18103.66	64530.36	1430.85	21877.14	

The companies include the largest passenger railway companies as well as companies with the most ridership. Some of these large companies deal only in infrastructure, hence the exclusion of some financial information. Other financial exclusions are due to the generality of some financial statements.

Fig. 3 presents the costs for the ten passenger railway companies and validates the results through a basic trend analysis. There is an observed correlation between highest cost and importance of the cost to a company. The observed trend in Fig. 3 identifies the top three (critical) costs for passenger railway companies. Companies seek to reduce expenditure to optimum thus capital debt, infrastructure and operations have earned priority for the purposes of investigation. It is argued that because maintenance has a contribution to the expenditure of infrastructure, it should also be investigated as a critical cost.

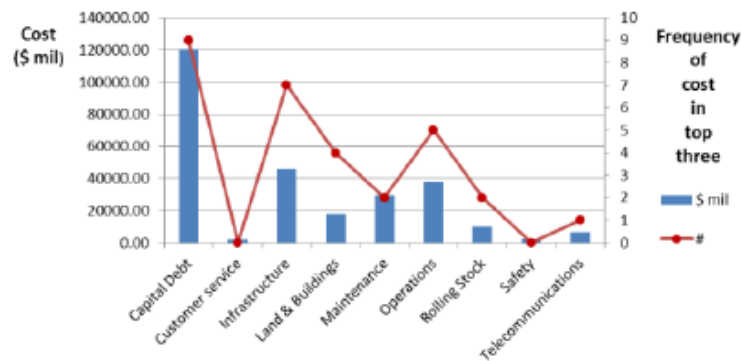


Figure 3: The total of each 'cost' across ten railway companies and the number of times each cost was a critical cost to an individual company.

An engineering investigation is to be conducted on critical company costs, thus capital debt can be ignored. Another concern is that companies spend too little on safety, considering that safety is one of the drivers on the mission statement of many of the investigated companies. Thus a common thread is sought in literature between infrastructure, operations, maintenance and safety. The proposed research idea is to improve infrastructure expenditure through better maintenance and operations tactics, thus improving safety as well.

3 Literature Study

Maintenance, operations and safety topics are investigated to see how they relate to the infrastructure division of a passenger railway service company.

3.1 Maintenance

The current established method for railway track maintenance in South Africa is condition-based maintenance, using a track geometry rail car [6]. Track cars are driven along the South African circuit to determine poor sections of track by measurement and statistical analysis. Standard deviation of track geometries from the design conditions are used to construct a Track Quality Index (TQI). Poor track sections are thus identified and maintenance schedules are created around this information.

New approaches in preventative maintenance on rail tracks are being investigated by Minsili et al. [7]. They have developed a ballast renewal strategy which allows for improved long-term health of the track network, thus saving replacement and maintenance cost. The effectiveness of this method could pave the way to a movement in the direction of preventative maintenance for tracks. Oyama & Miwa [8] developed an All-Integer Linear Programming (AILP) optimisation model for optimum railway track scheduling. The model minimises maintenance cost and maximises aggregate ride quality and safety levels of railway track. The schedule of a multiple tie tamper (MTT) is used for optimisation. This machine repacks track ballast and corrects deviant track geometry. Tamping is a condition type maintenance method, but advance planning and route optimisation provides structure for the addition of preventative maintenance methods. Higgins [9] considered a train operations schedule and minimised the number of times that scheduled maintenance should clash with train operation. A tabu local search optimisation method was used to process the large amount of decision variables in the problem. This shows that track maintenance not only needs to have optimal routes for maintenance but also specific times for maintenance, in accordance with a train operating schedule. An 8% reduction in interference delay was achieved on a train schedule.

Current track maintenance methods involve condition-based and reactive maintenance. Schedules are planned in advance but reliability and optimisation strategies are lacking. New research in preventative maintenance, coupled with condition-based approaches seems to be a rich avenue for exploration.

3.2 Operations

The infrastructure of a rail company affects operations when infrastructure faults cause delays and speed restrictions. Maintenance of infrastructure can also have schedule clashes with train operations, all of which have an opportunity cost. Capacity issues lead to capital expenditure for upgrading infrastructure. Performance indicators developed to trace railway operations can be used to monitor infrastructure performance.

The press in South Africa (official) reported that the Railway Safety Regulator issued an improvement initiative to Metrorail Western Cape, in March 2014 [10]. Speed restrictions of 15km/h were imposed on lines outside Cape Town station as the infrastructure was declared to be in unsafe condition. In 2014, Network Rail under-spent £1.2bn on maintenance [11]. Infrastructure faults caused a 1-5% increase in missed punctuality targets. This increase in delays caused a £53 million fine by the Office of Rail Regulation. Sudden failures cause a blockage in service as trains are delayed while faults are being fixed. This is the case for Network Rail. Poor maintenance of infrastructure ultimately leads to unreliability, discounting the top customer need according to the customer needs

analysis. These negative effects on operations could be better controlled with more structured track maintenance.

Railway infrastructure is linked directly to the capacity of a rail network, which is governed by size of the locomotive fleet, the extent of infrastructural development and train scheduling. During peak hours, the most trains will be in service and infrastructure failures could have an effect on train delays across the fleet. Scherbanin [12] investigated Russia's railway problems, highlighting capacity bottlenecks as one of the most significant problems. A capacity bottleneck could be seen as a single-carriageway train line, which is not only an operational complexity, but also a critical problem area when considering infrastructure failures. This leads to a conclusion that high traffic single-carriageway tracks have to be maintained to a higher standard, with a higher reliability than other track sections. Gevert [13] also encountered capacity issues during company expansion, during Brazil's Carajas railway project. Solution methods included doubling up carriageways, at a huge capital expense and also increasing train length. Each solution method requires higher maintenance expenditure on the tracks, thus capacity issues are a large driver of infrastructure expenditure increases. Infrastructure research should enable higher track reliability at a lower cost, thus alleviating capacity issues by a quantifiable amount. Trains will then be able to travel faster and more safely, with a lower risk of failure on bottleneck track sections.

Key Performance Indicators (KPI's) are often used in the operations division to monitor specific areas of performance and to compare and improve efficiency of processes. This tool can be applied to infrastructure to give helpful performance targets for maintenance activities. Ahren [14] provided a definition of a key performance indicator. A 'performance indicator is a measure capable of generating a quantified value to indicate the level of performance taking into account single or multiple aspects'. When researching improvement strategies for infrastructure, these performance indicators can be used as a guide-line to ensure that valuable improvements are being made. Banverket (Swedish rail) used Maintenance Performance Indicators (MPI's) that affect infrastructure, namely: train delays due to infrastructure, number of train disruptions due to infrastructure, capacity restrictions, markdowns in current standard (speed restrictions), total number of urgent inspection remarks and Track Quality Index [15]. These indicators are a guideline for research on improvement areas for infrastructure. Stenstrom et al. [16] used a link and effect model to convert railway business objectives in to KPI's that are then analysed and implemented. This method allows for useful captured data to be converted into output. This model was applied to the Iron Ore line in Sweden and it proved to be efficient, when used in conjunction with other computer analysis software. KPI's can be developed for infrastructure research for a case-study company. Research can then be applied to the company and the suggested improvements will be measured using KPI's. From these KPI's, an implementation strategy will commence to realise the discovered improvements.

3.3 Safety

Infrastructure failures have a safety impact on train passengers, with the most detrimental being derailment. Infrastructure related incidents also cause death and injury of company workers. Safety has been identified as the top priority for customers in the needs analysis presented and most railway companies analysed, declare safety as one of their key performance areas.

An analysis of fatal train incidents in Europe between 1980 and 2009 was conducted by Evans [17]. For the nine countries analysed, infrastructure was the second highest cause of fatal train collisions for seven countries and it was the fourth highest cause for two of the countries. There were a total of 277 fatal train collisions during this period. Kyriakidis et al. [18] identified infrastructure technical failures as a key precursor to railway accidents. They developed a methodology which seeks to reduce precursors such that accidents can be prevented. The methodology was applied to eighteen of the world's major metros. Investments in infrastructure to reduce risk of accident was agreed to be one of the solutions to mitigate precursors. Reliability improvement of railway was also listed as a pro-active solution in the infrastructure domain. American rail fatalities and incidents were investigated by Liu [19], with specific focus on infrastructure failures. The top cause of derailment was broken rail and welds, followed by track geometry defects. These failure modes not only cause a safety hazard, but also disrupt operation of the trains, causing further losses. Infrastructure failures caused more derailments than rolling stock, indicating that infrastructure is a high safety priority for railway companies. It was also proven that derailment risk decreases as the condition of the track improves. This indicates that there is a relationship between probability of failure and track condition, which opens up an interesting avenue of research for infrastructure reliability.

Thus, in order for passenger railway companies to keep a solid reputation, as well as maintain reliability standards and safe-guard human life, infrastructure improvement and maintenance should be taken seriously.

4 Solution methods through research

The researcher analysed literature on solution methods to the identified critical 'cost' problems. The idea is to identify solution methods even if they haven't been used for railway before. A research avenue is thus created in which problems can be tackled in more detail. Two potential solutions are discussed for further research consideration, such that practical implementation will arise.

The definition of reliability is provided as a guideline for discussed methods: 'Reliability is the probability that a component will operate correctly for a specified portion of time (design-life) under the design operating conditions (amp, temperature, force) without failure [20].'

4.1 TQI derailment solution

The Track Quality Index (TQI) is a current tool used in the railway industry to monitor the condition of the tracks and thus schedule maintenance based on track condition. This is a useful maintenance tool, but it lacks the power of scientific methods such as reliability. Thus, a relationship is sought between TQI and reliability to capture the ease of using TQI, with the powerful output of reliability methods. With TQI, a standard deviation statistic is assigned to a track section based on the deviation from design geometric track parameters.

A reliability block diagram can be developed for a track system based on reliability methods for each track component. The basis of this reliability model is the probability of train derailment caused by geometric rail irregularities. Derailment is considered to be failure of the track system according to the definition, thus a reliability basis for the system is established. This method produces a single reliability statistic for each track section, based on estimation, from which a track maintenance schedule can be built. Critical reliability statistics require more immediate maintenance attention.

For this solution method to succeed, reliability methods for each track component need to be identified from literature. These methods each need to be applied to a case study which has failure information for each track component for an entire rail network. Each rail section also needs to have up-to-date geometric measurements, which are available. As TQI is based on standard deviation of rail and the rail reliability also depends on standard deviation, some relationship can be constructed between TQI and rail reliability. TQI derailment is a heavily computational method and it requires extensive data extraction. It also requires efficient capturing of data by the company, which may not be the case in third world countries. Data this extensive may also not be available by most railway service companies.

4.2 Track failure reliability map

As a starting point, a fault tree analysis and a Failure Modes and Effects Criticality Analysis (FMECA) must be directed in order to understand the cause and effect of railway track failures. The failure modes must be well understood as only failure modes that affect the operation of trains will affect track reliability. This is according to the definition, as any component acting apart from design conditions (failure) is considered to be unreliable.

Once the failure modes are understood, a dataset of track incidents must be obtained. This dataset is then filtered to only include failure modes that affect the operation of trains. A speed restriction or potential train stoppage is considered to be an effect worthy of unreliability. Reliability methods can be applied to the remaining data to determine a reliability map for the train network. Each track section should have a reliability statistic, which will allow for an optimisation of maintenance routes, required to restore track sections to a more reliable state.

The reliability statistics will also give train drivers an idea of how to navigate the track sections ahead, when embarking on a trip. This reliability map will not only reduce wear and tear of rolling stock components through better route navigation, but it will also save maintenance cost of infrastructure through maintenance route optimisation.

4.3 Discussion of solution methods

The two discussed solution methods each compute reliability from a different reference point. The TQI derailment solution uses derailment as a baseline reliability failure. Track failure reliability map views the change in operating conditions in the rail network due to infrastructure failures as unreliability's in the system. The two different approaches discussed are relevant because each have different degrees of applicability to a practical context. The changes in operating condition approach uses actual operations and failure data to arrive at network reliability, rather than using more abstract TQI information. Probability of derailment is the most theoretical as some train networks have very few derailments in actuality. This method goes by the assumption that railway conditions are driven by safety, away from the fears of derailment. Therefore, in applying theory to practice, the 'Track failure reliability map' is the most relevant method.

The difficulty of data capture also needs to be low to ensure that the chosen method can be applied to different railway companies with relative ease. The Track failure reliability map solution has the data capture method with the easiest access to information that most companies are likely to have available. TQI derailment, on the other hand is complex as a vast array of individual component failures need to be available to produce reliability information. In practice, such detailed information may not be available. The track failure reliability map is thus the preferred method for data capture purposes.

The final judgement of the solution methods is the computational difficulty of the proposed solution. TQI derailment has the most complex computational process as it contains a number of reliability method computations for each railway component in a track section. Track failure reliability map has the least complex computational process as reliability statistics simply need to be computed from spread out actual component failures, which are grouped into separate track sections. Each method could use maintenance schedule optimisation thus computation time for this procedure is effectively even.

5 Conclusion

The 'Track failure reliability map' solution clearly has the most benefits as a research initiative. It is difficult to say which method would reduce company costs the most, but it certainly wouldn't be the best method if it wasn't possible to implement. Track failure reliability map is therefore the best solution to

eliminate critical problems from the desk of CEO's in the passenger railway industry. It is further noted that the research objectives were met through this solution method.

6 References

- [1] Maluleke, K., Instilling safety culture in the passenger rail transport industry within the South African context. *Journal of transport and supply chain management*, 7(1), 2013.
- [2] Fogel, G., It's time to start having the right conversations about ISO 55000. *Uptime – for maintenance reliability and asset management professionals*, (7), 2013.
- [3] Takai, S. & Ishii, K., A Use of Subjective Clustering to Support Affinity Diagram Results in Customer Needs Analysis. *Concurrent engineering, research and applications*, 18(2), pp. 101–109, 2010.
- [4] Blanchard, B. S. & Fabrycky, W. J., *Systems Engineering and Analysis*, Fifth edition, Pearson Education: New Jearsey, 2011.
- [5] Marquez, F., Lewis, R., Tobias, A., Roberts, C., Life cycle costs for railway condition monitoring. *Transportation Research Part E*, 44, pp. 1175–1187, 2008.
- [6] Zaayman, L., Track condition monitoring and analysis for effective maintenance planning. *Civil Engineering: Magazine of the South African Institution of Civil Engineering*, 19(4), 2011.
- [7] Minsili, L. S., Jeremie, M. D., Tsebo Simo, G. L., Simo, C., Preventive maintenance of railway tracks: Ballast performance anticipation in the Cameroon railway. *Research Journal of Applied Sciences, Engineering and Technology*, 4(5), pp. 398–406, 2012.
- [8] Oyama, T. & Miwa, M., Mathematical modelling analyses for obtaining an optimal railway track maintenance schedule. *Journal of Industrial and Applied Mathematics*, 23(2), pp. 207–224, 2006.
- [9] Higgins, A., Scheduling of Railway Track Maintenance Activities and Crews. *The Journal of the Operational Research Society*, 49(10), pp. 1026–1033, 1998.
- [10] Railway Safety Regulator. RSR media release. Online.
<http://www.engineeringnews.co.za/article/rsr-finds-prasa-transnet-w-cape-rail-lines-to-be-in-poor-condition-2014-03-28>
- [11] Topham, G., Network Rail fines more than £50m for late trains. Online.
<http://www.theguardian.com/business/2014/jul/07/network-rail-fined-50m-pounds-late-trains>
- [12] Shcherbanin, Y. A., Some problems of Russia's railway infrastructure. *Studies on Russian Economic Development*, 23(1), pp 37–47, 2012.
- [13] Gevert, T., Carajás rises to capacity challenge. *International Railway Journal*, 47(6), pp. 32–33, 2007.
- [14] Ahren, T., A Study of Maintenance Performance Indicators for the Swedish Railroad System. Licentiate thesis - Luleå University of Technology, 47, 2005.

- [15]Ahren, T. & Parida, A., Maintenance performance indicators (MPIs) for benchmarking the railway infrastructure: A case study. *Benchmarking: an international journal*, **16**(2), pp. 247–258, 2009.
- [16]Stenström, C., Parida, A., Galar, D., Kumar, U., Link and effect model for performance improvement of railway infrastructure. *Journal of rail and rapid transport*, **227**(4), pp. 392–402, 2013.
- [17]Evans, A., Fatal train accidents on Europe's railways: 1980-2009. *Accident analysis and prevention*, **43**(1), pp. 391–401, 2011.
- [18]Kyriakidis, M., Hirsch, R., Majumdar, A., Metro railway safety: An analysis of accident precursors. *Safety Science*, **50**(7), pp. 1535–1548, 2012.
- [19]Liu, X., Barkan, C. P., Rapik Saat, M., Analysis of Derailments by Accident Cause: Evaluating Railroad Track Upgrades to Reduce Transportation Risk. *Transportation research record*, **2261**, pp. 178-185, 2011.
- [20]Elsayed, A. E., *Reliability Engineering*, 2nd edition, John Wiley & Sons, 2012.

Appendix G

South African Journal of Industrial Engineering Article

In March 2015, Mark attended the CIRP 25th Design Conference in Haifa, Israel. He presented his work on reliability centred perway maintenance tools and received useful input from members of the Technion Haifa, as well as international attendees. Among those presenting at the conference was The vice-president of research for General Motors and a member of the engineering team from the 'Iron Dome' missile defence project. Mark's article was submitted to the South African Journal of Industrial Engineering and was accepted. The article is in the pre-publication editing process.

USING TQI TO QUANTIFY THE RELIABILITY OF A SECTION OF PERWAY

M.D. Mc Naught, N. Fourie¹

¹Post-graduate Department of Industrial Engineering, University of Stellenbosch

ABSTRACT

The railway industry in South Africa is in need of constant improvement, especially if road traffic can be reduced by increasingly proficient rail transit. To save maintenance costs and to make train transit safer, maintenance tools need to be adapted from a 'condition' focus to a reliability focus. The 'condition' tool TQI (Track Quality Index) is investigated in this paper. A relationship is sought between TQI and rail reliability by considering scientific reliability methods developed in literature. The definition of reliability is applied to the rail environment and is coupled with literature investigations to develop a reliability block diagram for a section of perway, which is a network of components on a rail track. A scientific basis for the maintenance of perway components is thus established. All structural and mechanical components in a rail network should eventually have associated reliability statistics.

OPSOMMING

In Suid Africa benodig die spoor bedryf voortdurende verbetering, veral as spoor vervoer kan bydra tot die vermindering van padverkeer. Om instandhoudingskoste te spaar en spoorvervoer veiliger te maak, moet instandhoudings tegnieke se fokus verander word van 'n 'kondisie' na betroubaarheid. Die SKI 'kondisie' indeks (Spoor Kwaliteits Indeks) word ondersoek en die verband tussen SKI en spoor betroubaarheid gesoek, deur wetenskaplike tegnieke in die literatuur te ondersoek. Die definisie van betroubaarheid word aangewend in die spoor omgewing gebaseer op die resultate van die literatuurstudie, en 'n betroubaarheid blok diagram word opgestel vir 'n spoor seksie, wat 'n netwerk van komponente op die spoor insluit. 'n Wetenskaplike benadering word dus toegepas op die spoorbaan komponente, en alle meganiese komponente in 'n spoor netwerk behoort uiteindelik met betroubaarheids statistieke geassosieer te word.

1. INTRODUCTION

Determining the reliability of complex mechanical systems is no exact science. Large systems such as power generation cycles, trains and valve networks need to be reliable to be economical. Large capital investments exist for these projects and safety is of critical importance for the operation of these systems. Scoping in on the mechanical component network of a train and supporting structures, it is perceived that maintenance accounts for a large portion of the total expenditure in the life of the system [1]. Internationally, railway companies and governments are investing in reliability research of railway systems to save cost on maintenance and improve the safety of systems.

1.1 Application of reliability research

Railway companies in South Africa such as Transnet and PRASA conduct applied research to quantify the reliability of mechanical component networks. The PRASA research group is interested in using practical maintenance tools to assess the condition of systems which hopefully informs the reliability of such systems. One such tool to be investigated is TQI (Track Quality Index) which is a threshold indicator of track condition, based on rail geometric irregularities. A scientific definition of TQI is presented in section 2. This article investigates the TQI to quantify the reliability of a section of perway. A section of perway is defined as the rail, rail clip, rail pad, sleeper and ballast assembly for a 200 meter 'break section' (standard section) of track. A diagram of the subsystems of a section of perway is presented in Fig 1.

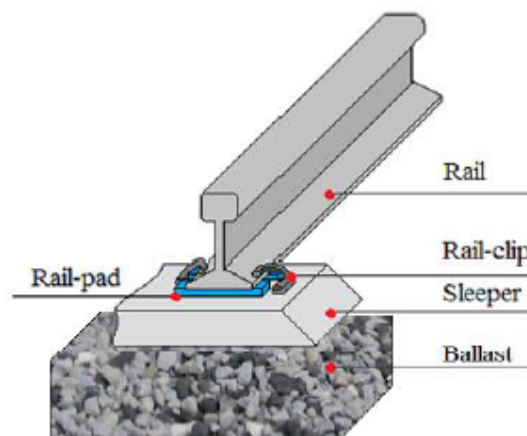


Figure 1: Simplified rail profile for a section of perway

1.2 Research questions and objectives

TQI is directly related to the railway track which is a component of a section of perway. The reliability of a section of perway can begin to be quantified if a relationship can be developed between TQI and track reliability. This relationship must then be used to inform a reliability model of the perway.

From this reasoning, the following research questions are formulated:

- What is the relationship between TQI and rail reliability?
- What is the influence of rail reliability on the perway reliability network?

From the research questions, the following objectives are set:

- Show that a relationship exists between TQI and rail reliability.
- Determine whether TQI is a useful input to determine rail reliability.
- Develop a reliability block diagram of the perway component network.

2. DEFINITIONS OF WORKING METHODS

Prior to investigating academic literature, definitions of key concepts are explored to familiarise the reader with the working units of the investigation.

2.1 TQI

The geometry of a railway track changes due to rail wheel contact, sleeper failure and ballast shift. Changes in the track geometry result in track irregularity. Track irregularity is defined by five geometry parameters, namely: average vertical alignment (PRA), average horizontal alignment (ALA), twist (TWT), super elevation or cross level (SUP) and track gauge (GAU) [2]. Mean value measurement of track irregularity is conducted on 200 meter break lengths of track by the Plasser IM2000 recording car, instituted by Transnet. The track irregularity of a break length is quantified by Track Quality Index (TQI), which is the sum of standard deviations of the five track irregularity parameters. The South African method for TQI as discussed is simple in comparison to that used internationally, such as the TGI index used in India [3], which has weighted components. China's system is also more extensive than in South Africa, with 7 geometric parameters considered [4]. Xu et al. [4] develop an equation for TQI which is consistent with measuring techniques used by Transnet and PRASA.

$$\begin{aligned}\sigma_i &= \sqrt{\frac{1}{n} \sum_{j=1}^n (x_{ij}^2 - \bar{x}_i^2)} \dots [4] \\ TQI &= \sum_{i=1}^5 \sigma_i \dots [4] \\ \bar{x}_i &= \frac{1}{n} \sum_{j=1}^n x_{ij} \dots [4]\end{aligned}\tag{1}$$

Where σ_i is the standard deviation of measurements of the i_{th} track irregularity parameter at mileage points in the break length, x_i is the average of measurements of the j_{th} track irregularity parameter, x_{ij} is the measurement of the i_{th} track irregularity parameter at the j_{th} mileage point in the break length and n is the number of measurement points in the break length.

A maximum TQI of 7.5 would be allowed for high speed passenger lines and heavy haul lines, which are the operating infrastructure for PRASA. TQI therefore is not concerned with a severity scale but rather applies a threshold value that would be broken if the track geometric deviations are large [2].

2.2 Reliability

‘Reliability is the probability that a product will operate properly for a specified period of time (design life) under the design operating conditions (temperature, load, volt) without failure. In other words, reliability may be used as a measure of the system's success in providing its function properly during its design life [5]. ‘

The reliability of the perway depends on the reliability of the mechanical components that make up the perway. Using the definition of reliability, each mechanical component (or sub-system) must have a quantifiable measure for reliability. This is achieved by comparing the definition of reliability to the life cycle of the mechanical component in question. Thus, each mechanical component must meet the definition checklist:

- Probability component
- Time component
- Operating conditions
- Failure mode

As TQI relates directly to the rail pair subsystem, the definition is used to help clarify the relationship between TQI and rail reliability. Thus it will be shown that TQI is a function of perway reliability.

Although each type of component in the perway system has an individual definition for reliability, the general reliability definition for each component, according to the researcher, is: 'a reliable component of the perway system is one that enables the train to run on the track without derailment or to a lesser extent, maintains the stated operating conditions of the train.'

3. RESEARCH METHODOLOGY

As the reader now has an understanding of the problem case and the definitions applied in the research, the methodology is presented for the research in Fig. 2.

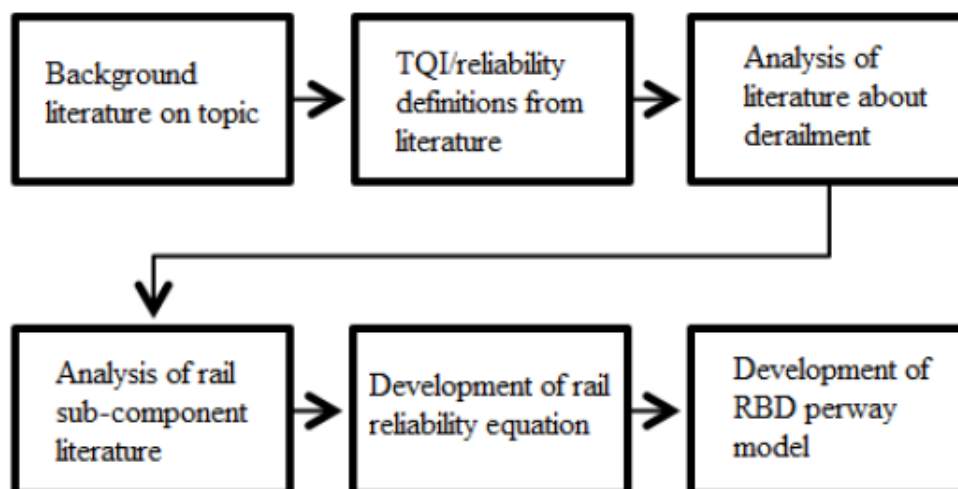


Figure 2: Research methodology

Through the definition of reliability, a link will be constructed between TQI and reliability. Reliability methods will then be investigated to develop a reliability method that can be used to determine the reliability of a rail. A 'bridge' shall then be constructed between TQI and a rail reliability method to evaluate the usefulness of TQI to determine rail reliability. Although it is possible that rail reliability could be determined without TQI, the practicality and simplicity in determining TQI is a desirable quality that would make reliability calculation easy if a supporting method can be found. A perway reliability model is the next step in the process.

4. LITERATURE INVESTIGATION

Literature from various sources was analysed to provide methods to determine the reliability of the mechanical components in a perway network. Specific emphasis was placed on a rail reliability method which relates to TQI. The objective of TQI vs. reliability relationship clarification is extended to the time domain at the end of the chapter.

4.1 Reliability of rail

The reliability of the rail pair on a railway track is quantified using the definition checklist. The failure mode of the track is the geometric condition of the track such that derailment will occur with 100% probability. The time component is the design life of the rail and the operating conditions are those specified by the South African rail standard. The probability of derailment (P_f) due to rail geometric irregularities is used as the failure mode of the track pair.

$$r_{rail_pair} = 1 - P_{f_{geom}} P_{f_{profile}} \quad (2)$$

The self-explanatory equation (3) thus quantifies the reliability of the rail pair based on series reliability system theory. Note that a break length of 200 meters is considered for reliability calculations (South African standard).

4.1.1 Probability of derailment due to geometric parameter irregularities

Mohammadzadeh et al. [6] used five geometric input parameters with random values to determine derailment probability using a Monte Carlo method. The Nadal criterion for derailment was considered in the formulation of the probability of failure. Nadal developed a derailment coefficient limit, L/V , which is a limit on the lateral over the horizontal forces experienced by a train wheel in operation. The equation used to calculate derailment probability is described. Further detail on the method is presented in the literature.

$$P_f = E \left(I(v) \frac{f_x(v)}{h_v(v)} \right) \approx \bar{P}_f = \frac{1}{N} \sum_{j=1}^N \left(I(v_i) \frac{f_x(v_i)}{h_v(v_i)} \right) \dots [6] \quad (3)$$

Where: $h_v(v_i)$ is the joint probability function of density sampling

$f_x(v_i)$ is the main joint density function of random variable

N is the total number of tests for Monte Carlo analysis.

$I(v_i)$ is the indicator function with a value of one if x is located in failure region
and zero if x is located in safe region.

Further detail on the first term of the above equation is described in the literature.

The methodology used in the literature can thus be followed to determine probability of failure based on track irregularity parameters. An input measurement from a recording rail car is necessary to use this method. The final step is the use of probability equation (3). The effect of non-critical perway components (rail pad failure and ballast distribution) on the probability of failure due to geometric input parameters is of importance to the accuracy of r_{rail_pair} . As rail pad failure and ballast distribution are some of the inputs that generate track geometry irregularities, their effect can already be accounted for as part of the random nature of track irregularity parameters.

4.1.2 Probability of derailment due to rail profile wear

Mohammadzadeh et al. [7] have developed a methodology for the determination of train derailment probability which incorporates the use of a track recording car. The track recording car measures the rail profile using a non-contact method. The rail profile can also be measured by the IM2000 track recording car used by PRASA. Van der Merwe [8] confirms the performance capabilities of the IM2000. The methodology for the calculation of probability of derailment due to rail profile wear is presented in Fig. 3. Further detail on the method and parameter definitions is presented in the literature.

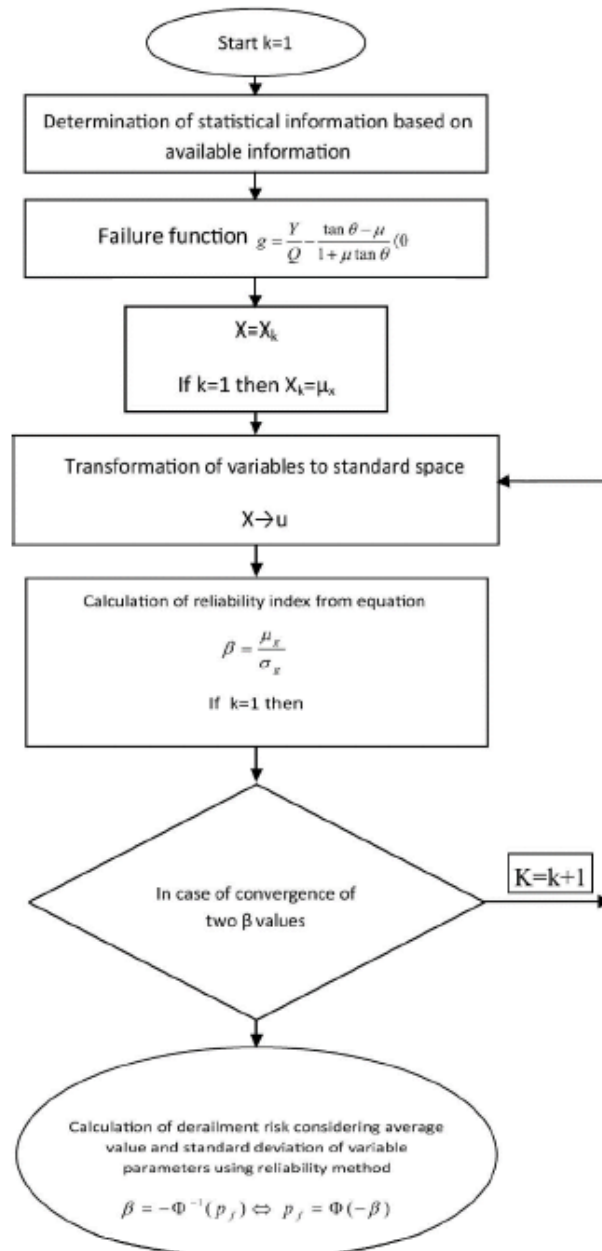


Figure 3: Methodology to determine train derailment probability [3]

4.2 Reliability of sleeper

Zhao, et al. [9] discusses reliability analysis for railway sleepers using track condition information. He develops a reliability model for a cluster of sleepers with consecutive failed sleepers, using k-out-of-n principles. The failure mode for a sleeper considered here is the brittle fracture of the sleeper such that it can no longer support anticipated loads according to specification. The reliability for an individual sleeper as well as four cases for sleeper cluster reliability are considered. The reliability of k=2 and k=3 sleeper systems was developed from cluster reliabilities [9]. The reliability equation for individual sleeper failure is extracted by the researcher for introductory purposes. The article can be

referenced for reliability calculations for the four cases discussed and for in-depth system reliability calculations.

A Weibull distribution could possibly be used to describe the distribution of the lifetimes of sleepers. The raw data for sleeper failure can be collected from the PRASA infrastructure database. The reliability of an individual sleeper at time t is presented as

$$r(t) = \exp[-(t/\beta)^\alpha], \quad \alpha, \beta > 0. \quad [9] \quad (4)$$

Where α and β are shape and scale parameters respectively. These parameters are determined by the analysis of existing sleeper failure data.

The reliability, r_t , of a single sleeper at time $t_0 + \tau$ and its probability of failure, q , given that it is functional after inspection at time t_0 may be given by

$$r_t = r(t_0 + \tau)/r(t_0) \quad [9] \quad (5)$$

$$q = [r(t_0) - r(t_0 + \tau)]/r(t_0) \quad [9] \quad (6)$$

Equation (5) allows a sleeper reliability prediction for a future time. This will allow the reliability block diagram to be developed to have future reliability predictions, so long as future reliabilities can be calculated for other components. It is further noted that because a distribution of failures is used to determine individual sleeper reliability; the effect of the rail, rail pad and ballast condition on sleeper reliability has been accounted for by the nature of these components interacting with a sleeper, adding to the cause of failure.

4.3 Reliability of ballast

Nurmikolu [10] conducted condition assessment of ballast and substructure and identified GPR (Ground Penetrating Radar) as a real-time tool to monitor ballast condition. From this research, a failure mode for ballast can be determined and over time, a statistical basis can be developed for ballast failure. Ballast failure can be defined as the percentage ballast material that passes through a specific sieve opening size that is more than specification. According to Silvast, et al. [11], a GPR fouling index exists for 5 meter track sections based on real-time signal data. This data can be averaged for 200 meter track sections and can be monitored on a real track over time. A reliability statistic is generated for this data by fitting observed failures to a known distribution. Nurmikolu [10] suggests taking samples from beneath the sleeper edges as this is where degradation is most concentrated. This provides a conservative reliability estimate. Sadeghi, et al. [12] gives equal weighting between rail, sleeper, rail clip and ballast structural condition for maintenance, thus in a reliability block diagram, the ballast reliability measure satisfies equality.

Prescott [13] uses Petri Net modelling to capture the effect of ballast maintenance on rail geometric parameters. Thus based on track geometry measures, ballast maintenance is conducted. This method does not clarify the direct causes of track geometric irregularities, which could be due to subgrade instability, thus the direct method chosen is preferred. Al-Qadi [14] uses GPR methods to detect ballast fouling, which solidifies the argument that GPR is the future technology for ballast maintenance.

4.4 Reliability of rail pad

Rail-pad failure cannot directly cause train derailment or whole system failure. Rail-pad failure can contribute to rail clip failure as the rail pad decreases impact rail deflections, which cause wear and tear on rail clips. The rail pad failure can be characterised as the deterioration of the dynamic characteristics beyond the point of commercial specification. This deterioration can be attributed to fatigue which increases with number of train load cycles. Remennikov et al. [15] test the dynamic characteristics (stiffness and damping) of rail pads in a laboratory environment using a direct testing method. The tests were

conducted for rail pads after varying the ton hours of loading. As passenger rail has differing numbers of passengers, the lifetime of installed rail pads should instead be used to develop a relationship between lifetime and dynamic load for rail pads. Spot tests of rail pads would need to be conducted for each break section to account for the effect of local conditions on the deterioration of dynamic properties. Infrastructure databases will have the installation time for each rail installation and the lifetimes could be calculated from this.

Chang et al. [16] uses an accelerated loading method to evaluate the change in rail pad properties. Heat cycles were applied to the rail pads with 0~95 kN loadings. Rail pad thicknesses decreased with load frequency and pad displacement increased for a specific load case. The elasticity of the pads also increased. Arrhenius curves were fit to estimate rail pad degradation with time. Although this method has validity, the real test case as presented by Remennikov et al. [15] is preferred over the theoretical model [16].

4.5 Reliability of rail clip

The reliability of an individual rail clip is based on its contribution towards preventing a train from derailling. The failure mode of an individual clip would be the fracture or absence of the clip from its installed position. The choice of the selected failure mode is further solidified by Prasad & Srikan [17], who show failure modes for the South African Pandrol e-clip under simulated load cases.

The raw statistical failure data used to determine reliability of individual clips must be categorised as either belonging to a curved track break length or a straight track (tangent track) break length as rail clips in each of these sections experience different loading conditions. If these sections were lumped together, reliability estimates for failed rail clips would gain an error as rail clips on curves are expected to receive more wear-and-tear than those on tangent tracks. For a tangent track, train speeds are higher than curved tracks and tangent tracks do not experience horizontal force variations as significant as those on curved tracks. Marquis et al. [18] state that inward cant angles such as that of a curved track are more likely to fail truck side derailment criteria than the zero cant of straight sections as the derailment coefficient for an inward cant curved section is higher.

A distribution such as Weibull is used to determine the reliability of an individual clip. The distribution will be determined, based on the raw failure data observed. In contest to the chosen method, Mohammadzadeh et al. [19] pose a fatigue life reliability method for the determination of individual rail clip reliability. This method applies rain-flow method and Palmgren-Miner linear damage rule for crack nucleation life and Monte Carlo simulation for a first order reliability method estimation. The approach by Mohammadzadeh is weakened by its deviation from statistical basis and by the dynamic testing required to determine force vs. displacement behaviour of clips. As a result, this method cannot be easily reciprocated on a real-life, large scale system.

5. SCRUTINY OF RAIL LITERATURE

Literature employed and methods used to develop the reliability of the rail is scrutinised for practical applicability. The output of these methods is the practical use of TQI to allow reliability block estimates for rail reliability.

5.1 Probability of derailment due to track irregularity parameters

Mohammadzadeh et al. [6] use advanced simulation techniques to model the interaction between a train car and random tracks. A level three reliability technique is then used to determine the derailment probability due to track irregularity parameters. The methodology demands input parameters from a track measuring car, thus the method is practically applicable if a track recording car is available. Random tracks are generated

using a Fourier series function method and statistical analysis of the geometric input parameters. These methods can be practically applied to valid geometric parameter database information. The response surface method is then applied to approximate the limit state function. The importance sampling method then uses the limit state function to determine derailment probability. This is a powerful mathematic method which is practically applicable to a railway environment. Simulations take under ten minutes where more than 10000 track samples of 200 meter break length are used. Thus, the reliability of each rail pair break length in a large railway network can be easily computed. The method is very applicable and should be implemented in the railway industry.

5.2 Probability of derailment due to rail profile wear

Mohammadzadeh et al. [7] develops a complex scientific method for the determination of derailment due to rail profile wear. The validity of the methodology presented in Fig. 2 is evaluated. For application of the methodology to a practical working environment, historical rail geometric data needs to be available from track recording car measurements (for 200 meter break sections). Train speed and axle load information also needs to be available for specific locations on the railway network in application. In practise, the implementation of this method would require an accurate database repository. If this indeed exists, the method can be applied to the railway network with relative ease. Common scientific derailment theory is used in the paper, as well as statistical information relevant to the specific rail network. An improvement of the method would be to output a reliability statistics that range from one to zero, rather than using a reliability index. This would make reliability statistics more compatible with other components in a larger network. The probability of derailment is thus extracted as the output from the research. Each track section probability of derailment must be computed individually which requires ample computing time. With leading industry computer packages, this method could be used to determine the reliability of rail pairs.

6. TQI VS RAIL RELIABILITY OBJECTIVES

The use of TQI for the calculation of rail reliability is that it provides geometry input parameters for the rail. This is done by measuring geometry parameters from the Plasser rail car and following two different avenues to calculate TQI and rail probability of failure due to track geometric irregularities, respectively. Fig. 4 presents how the TQI is broken down into components which are inputs to the probability of derailment equations (due to geometric irregularities). Fig. 4 proves that TQI is a function of σ_i and P_f is a function of σ_i therefore a relationship must exist between TQI and P_f , as suggested in the research objectives.

As TQI exists as an indirect input to the chosen reliability method and little manipulation of the TQI is necessary to obtain direct inputs into the reliability method from TQI, TQI is deemed by the researcher to be a useful input to determine rail reliability. The second research objective is thus satisfied.

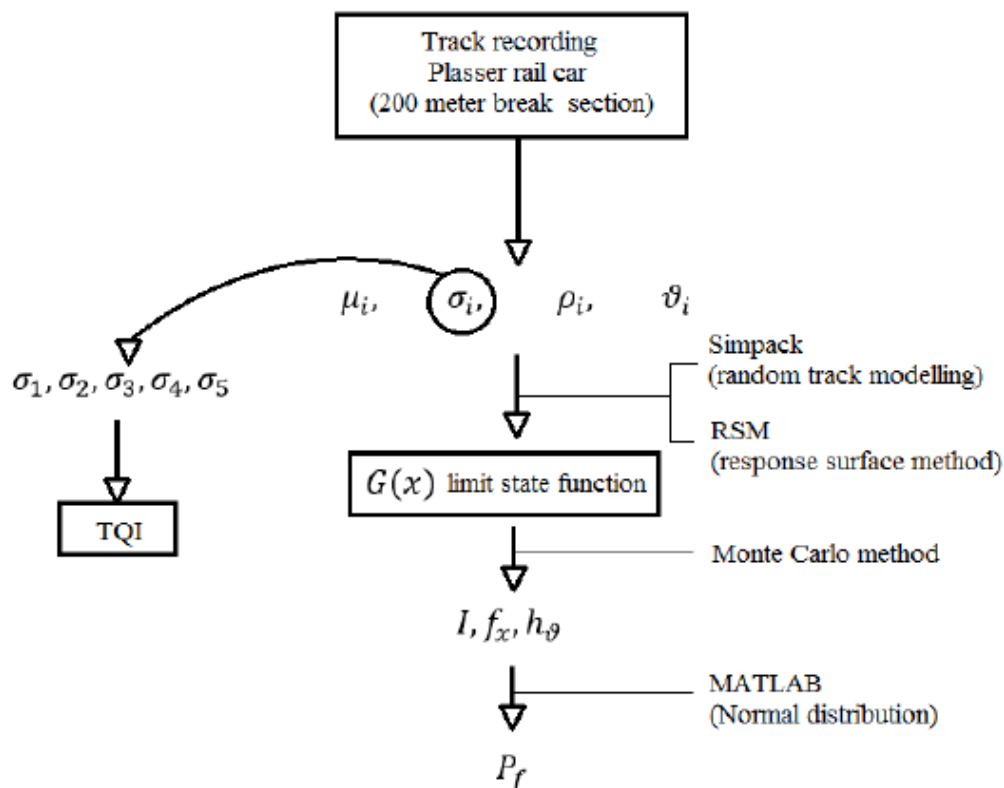


Figure 4: Relationship between TQI and rail derailment probability due to track irregularity.

6.1 Extension of TQI vs. reliability relationship

To satisfy the objective: 'show that a relationship exists between TQI and rail reliability' an instantaneous relationship of how reliability changes with TQI was determined. To improve upon this objective, a prediction of how TQI changes with time must be applied to this relationship such that a relationship of how reliability changes with time can be determined. The time changing nature of reliability is useful as maintenance plans are formed on the basis of future expected trends. Thus, the time changing nature of TQI is investigated.

Sadeghi [12] developed a track degradation model which determines how track condition indices change with time. The TGI (Track Geometric Index) similar to TQI was evaluated against time, train speed, TSI (Track Structural Index), loading and initial TGI. During tests on the Iranian rail line, all variables were fixed and only one of the discussed variables was allowed to change. The model ultimately produced a graph which presents the relationship between TGI and time. TGI and TQI have the same inputs but the calculation of each parameter is slightly different. The same methodology to produce this model can be applied to a South African rail context to produce a model for changes in TQI over time. An average estimate for the changing TQI for different categories of tested track sections is presented in Fig. 5. Note that the decreasing TGI trend (negative) is opposite from expected when considering TQI. This is because with TGI, the geometry irregularity parameters are inputted into a function with a negative exponential before summing together for the TGI. Detailed definitions can be reviewed in the literature.

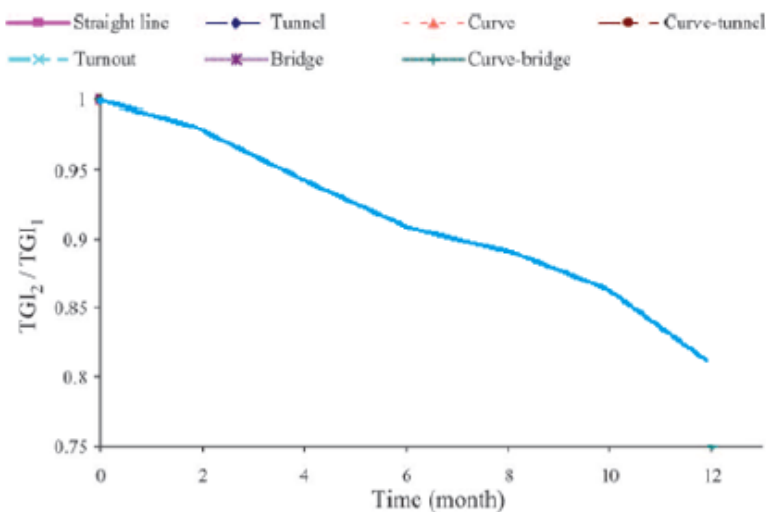


Figure 5: Averaged estimate for change in TQI between maintenance intervals (Adapted [6])

Xu [4] investigated a short-range prediction model for track quality index. A track irregularity prediction technique, SRPM-TQI was used in accordance with historical TQI calculated from waveform data. This approach applies the least squares method (LSM) to develop a prediction model for each break length of track. LSM produced errors of less than 8% for sixty break lengths of track. Abnormal deviations occurred when tamping or levelling operations were in progress during measurement intervals. Two methods were discussed to improve least squares predictions when these operations were in play.

6.2 Scrutiny of future TQI prediction methods

Sadeghi [6] and Xu [7] juxtapose two methods for future TQI prediction. Sadeghi [6] uses a scientific method of physical testing whereby all variables but one are fixed such that the relationship for changing a single variable can be determined. Xu [7], on the other hand, uses a statistical LSM to predict future TQI readings. The method by Xu [7] is more advantageous to calculate future TQI predictions than that used by Sadeghi [6] as the TQI method used by Xu [7] is the same method seen in the South African rail network. In addition, Xu [7] presents the advantage of testing predictions against an actual measurements (using error). Sadeghi [6] uses TGI degradation coefficients based on actual measurements but does not involve error nor validate the accuracy of variable testing. The LSM method by Xu [7] is thus the preferred short-range prediction model for TQI.

7. RESULTS

The RBD (Reliability Block Diagram) of a section of perway comprises of a network of components that each have their own reliability. If a component in series fails (reliability zero), then the system fails and if a component in parallel fails, the system as a whole can still operate without failure. The rail clip and sleeper systems have more conservative and less conservative options for failure. Rail pad and ballast systems are non-critical and their reliability blocks are independent from the perway system, being useful for maintenance purposes. The diagram is presented in Fig. 6.

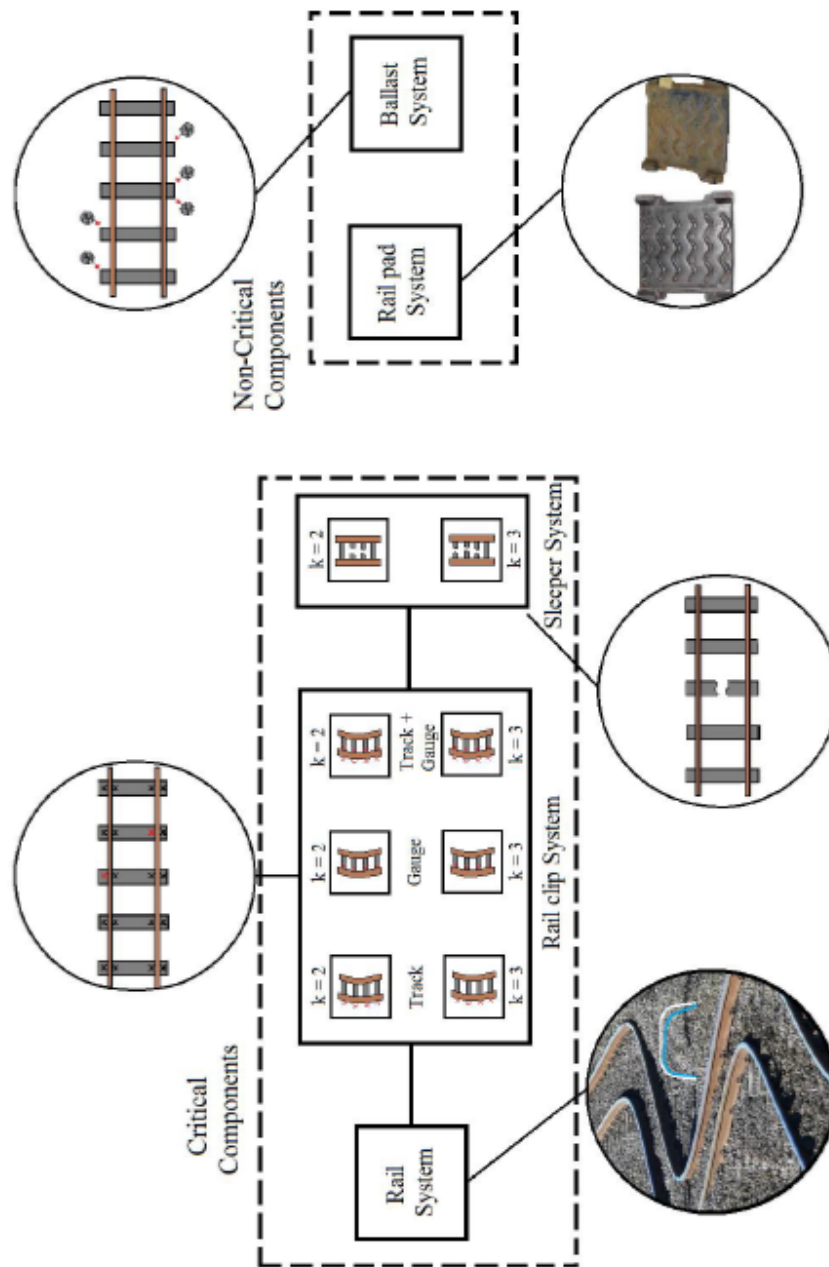


Figure 6: Reliability block diagram for a 200 meter section of perway, adapted [20], [15].

7.1 Reliability of the rail system

The rail system in Fig. 6 is considered to be a single unit for the purposes of reliability. The combination of track geometric irregularity and rail profile wear that causes the train to derail with 100% probability, is the failure mode of the rail pair. This would produce a reliability of zero in the RBD and total failure of the perway subsystem.

7.2 Reliability of the sleeper system

The sleeper system in Fig. 6 constitutes the series reliability block multiplication of $k=2$ and $k=3$ systems, presented in Fig. 7 [9]. The combination of these sleeper failures that causes the train to derail with 100% probability, is the failure mode of the sleeper block. This produces a 0 for sleeper system reliability in the RBD.

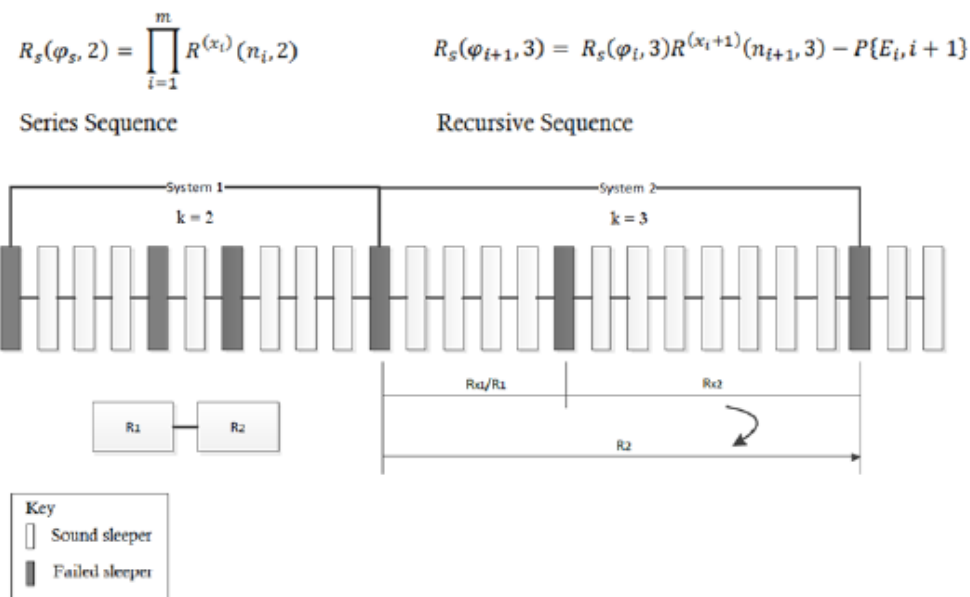


Figure 7: k -out-of- n reliability analysis for $k = 2$ and $k = 3$ systems [9]

7.3 Reliability of the ballast system

The ballast system, like the rail system in Fig. 6, is considered to be a single unit for the purposes of reliability. The ballast is a non-critical component in terms of derailment although a failure mode does exist against specification. The reliability block informs maintenance decision. The failure of the ballast to meet rail standard specification would produce a reliability of zero in the RBD and total failure of the perway subsystem.

7.4 Reliability of the rail pad system

The rail pad system in Fig. 6 is a system of multiple individual reliabilities, but because the rail pad is non-critical in terms of derailment, the reliability block is independent and the failure mode is identified for maintenance purposes. This mode is the failure of the rail pad to meet standard specification. This would result in a reliability statistic of zero for the reliability block.

7.5 Reliability of the rail clip system

The failure mode of the rail clip system would be the combination of individual clip failures that causes the train to derail. The derailment mechanism that rail clip failure can cause is called 'rolling track failure' [21]. This occurs when the lateral and vertical derailment forces from the rolling stock overcome the torsional stiffness and rail clip restraining forces and cause the track to pivot on its track side edge. The failure mode for a two dimensional section of track is presented in Fig. 8.

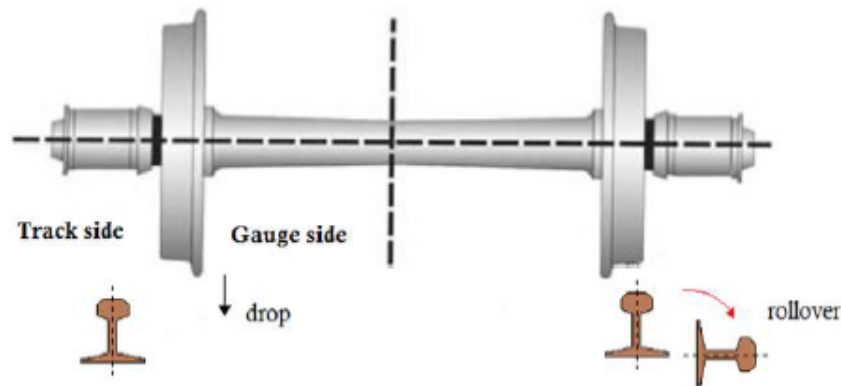


Figure 8: Rail rollover failure mode [21]

The rail rollover failure mode is further confirmed by Grøve et al. [22]. They measure pressure experienced by the rail seat for increasing coefficients of derailment. The findings reveal high pressure on the track side of the rail seat with little pressure on the gauge side for high coefficients of derailment. This indicates an outward rolling moment experienced by the rail.

For rail rollover failure modes, it is likely that a particular side of the of the rail experiences poor resistance to the moment forcing the rail over as continuous wear and tear has historically caused these types of failures, as discussed by Marquis [18]. The researcher developed a simplified beam model in Fig. 9 to illustrate the effect of clip supports and rail torsional stiffness on the rail rollover condition. This figure provides support to bolster the argument that continuous clip failure has a drastic effect on probability of derailment when compared to dispersed clip failures.

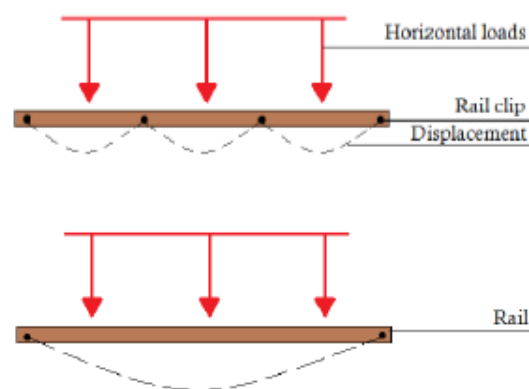


Figure 9: Top view of train loaded rail with rail clip supports

Fig. 9 shows that with consecutive failed supports, the rail is far more likely to bend. This bending acts as a catalyst for rail rolover as the horizontal forces are applied at the top end of the rail, initialising a tipping motion. This argument is supported by the US dept. of transportation [23] who show that for three consecutive clip failures, the gage widening resistance limit of a track section is exceeded, which is considered to be a system failure. The resistance to gage widening decreases as successive clip failures increase, according to the research.

Research on k-out-of-n sleeper failure by Zhao et al. [9] can be used to determine the reliability of a rail clip system. This research is consistent with observed reliability theory for consecutive k-out-of-n: F systems [5]. Six possibilities of failure are considered, due to the nature of consecutive clip failures and their effect on reliability: Two consecutive failures for track side clips, gauge side clips and a combination of track side and gauge side clips causing system failure. Also, three consecutive failures for track side clips, gauge side clips and a combination of track side and gauge side clips causing system failure are included in the six. Train derailment is most likely for consecutive failed track and gauge clips, less likely for gauge side consecutive failed clips and least likely for track side consecutive failed clips. The effect of each failure type can be observed by considering the moment equation around the rail edge for rolling rail derailment, equation (8). Fig. 10 presents a moment diagram of a rail section under loading from a train.

$$M = L'H - V'D - CX - CY \quad (8)$$

$$L' = L\cos\theta - V\sin\theta \quad (8a)$$

$$V' = V\cos\theta + L\sin\theta \quad (8b)$$

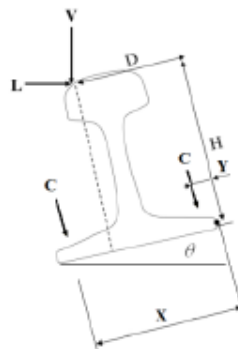


Figure 10: Moment diagram of rail section with train forces and clip forces

The researcher proposes that a sensitivity analysis be conducted to determine the effect of track side and gauge side rail clips on reliability. Fig. 10 indicates that track side clip forces aid the track torsional stiffness in resisting longitudinal and vertical forces less than gauge side clip forces as track side clips have smaller moment arms than gauge side clips. The variation of the L/V derailment coefficient and variation of D/H can be used for sensitivity purposes. From the sensitivity analysis, an estimate for a multiplier coefficient can be determined that will be applied to the reliability for consecutive track and gauge side failures respectively for $k = 2$ failure modes and $k = 3$ failure modes. Thus, if the same system is considered for varying potential failure modes, the reliability will be lowest for consecutive track side and gauge side failures, higher for track side failures and highest for consecutive gauge side failures.

The reliability block for the rail clip system presented in Fig. 6 is representative of the respective causes of derailment of the train and their associated reliability statistics,

considering a specific dataset of failed clips. A reliability of zero would be an extremely rare case.

8. CONCLUSIONS AND RECOMMENDATIONS

The original objectives are re-visited to draw a conclusion from the research paper. It was uncovered that TQI and track reliability are related as they both change according to the same input parameters. As TQI is a linear equation, it is easy to monitor the effect of changing input parameters on the TQI output. The effect of changing input parameters on rail reliability is not easily determinable. It is recommended that this is calculated by conducting a sensitivity analysis by changing only σ_i , the input parameter to the rail reliability equation. An equation can then be developed for rail reliability as a function of TQI.

The usefulness of TQI as an input to rail reliability is a debatable point. Without the input parameters for TQI, the rail reliability cannot at current be determined. If TQI information was available but the input parameters to TQI were not, then rail reliability could not be determined unless a strong relationship existed between TQI and rail reliability, pending a sensitivity analysis. Thus the most basic form of TQI (input parameters) is useful for the determination of rail reliability.

A reliability block diagram was developed which enables the calculation of perway reliability. This model can be extended to include a prediction model for the deterioration of reliability over time. The ability of this model to improve railway maintenance techniques is unquestionable. Moving away from condition based maintenance to reliability based maintenance is imperative for the cost reduction of maintenance and the improvement of safety in the railway industry.

9. REFERENCES

- [1] Esveld, C. 2001. Modern railway track. 2nd edition. The Netherlands: Delft University of Technology Publishing Service, 475-565.
- [2] Zaayman, L. 2013. *Mechanisation of track work in developing countries*. Independent publication, South Africa.
- [3] Sadeghi, J. Askarinejad, H. 2010. Development of improved railway track degradation models. *Structure and Infrastructure Engineering*, 6(6), pp 675-688.
- [4] Xu, P. Sun, Q. Liu, R and Wang, F. 2011. A short-range prediction model for track quality index. *Journal of rail and rapid transit*, 225(F3), pp 277 -285.
- [5] Elsaved, A. E. 2012. Reliability Engineering. 2nd edition. John Wiley & Sons. *Note reliability engineering and risk analysis used.
- [6] Mohammadzadeh, S. Sangtarashha, M and Molatefi, H. 2011. A novel method to estimate derailment probability due to track geometric irregularities using reliability techniques and advanced simulation methods. *Archive of applied mechanics*. 81(11), pp 1621-637.
- [7] Mohammadzadeh, S and Ghahremani, S. 2012. Estimation of train derailment probability using rail profile alterations. *Structure and infrastructure engineering*, 8(11), pp 1034-1053.
- [8] Van der Merwe, G. Zaayman, L.C. 2001. IM2000 Infrastructure measuring car: the application of recording results. *Rail Engineering International*, 4.
- [9] Zhao, J. Chan, A. Burrow, M. 2007. Reliability analysis and maintenance decision for railway sleepers using track condition information. *Journal of the Operational Research Society*, 58(8), pp 1047-1055.
- [10] Nurmikolu, A. 2012. Key aspects on the behaviour of the ballast and substructure of a modern railway track: research-based practical observations in Finland. *Journal of Zhejiang University A*, 13(11), pp 825-835.

- [11] Silvast, M. Nurmikolu, A. Wiljanen, B. Levomaki, M. 2010. An inspection of railway ballast quality using ground penetrating radar in Finland. *Journal of rail and rapid transit* 224(F5), pp 345-351.
- [12] Sadeghi, J. M. Askarinejad, H. 2011. Development of track condition assessment model based on visual inspection. *Structure and Infrastructure Engineering*, 7(12), pp 895-905.
- [13] Prescott, D. Andrews, J. 2013. A track ballast maintenance and inspection model for a rail network. *Journal of risk and reliability*, 227 (3), pp 251-266.
- [14] Al-Qadi, I, L. Xie, W. Roberts, R. 2008. Time-frequency approach for ground penetrating radar data analysis to assess railroad ballast condition. *Research in nondestructive evaluation*, 19(4), pp 219 -237.
- [15] Remennikov AM. Kaewunruen S. Ilkaunieks K. 2006. Deterioration of dynamic rail pad characteristics. *Proceedings of the conference of railway engineering*.
- [16] Chang, S, W. Hyun, S, P. 2014. Using lifetime prediction of rail pads for high speed train. *International Journal of Mechanical, Aerospace, Industrial and Mechatronics Engineering*, 8(4).
- [17] Prasad, C. Srikan, S. Functional failure analysis of a Pandrol clip. 2009. *International Journal of Applied Engineering Research*, 4(11), pp 2329.
- [18] Marquis, B, P. Muhlanger, M. Jeong, D, Y. 2011. Effect of wheel/rail loads on concrete tie stresses and rail rollover. *Proceedings of the ASME 2011 Rail Transportation Division Fall Technical Conference*.
- [19] Mohammadzadeh, S. Ahadi, S. Nouri, M. 2014. Stress-based fatigue reliability analysis of the rail fastening spring clip under traffic loads. *Latin American Journal of Solids and Structures*, 11(6).
- [20] Petley, D. 2010. The Canterbury Earthquake: Images of the distorted railway line [Online]
Available: <http://blogs.agu.org/landslideblog/2010/11/02/the-canterbury-earthquake-images-of-the-distorted-railway-line/> [22 September 2014].
- [21] Iwnicki, S. Ed. 2006. *Handbook of railway vehicle dynamics*. CRC press.
- [22] Greve, M, J. Rapp, C, T. Dersch, M, S. Edwards, J, R. Barkan, C, P, L. 2013. Quantification of Concrete Sleeper and Elastic Fastening System Demands Utilizing Concrete Sleeper Rail Seat Contact Area. *Proceedings from 2013 World Congress on Railway Research*.
- [23] US dept. of transportation. 2011. Effect of missing or broken fasteners on gage restraint of concrete ties. *Research results - FRA office of railroad policy and development*.

List of References

- Åhrén, T. (2005). *A study of maintenance performance indicators for the Swedish railroad system*. Ph.D. thesis, Citeseer.
- Al-Fawzan, M.A. (2000). Methods for estimating the parameters of the Weibull distribution. *King Abdulaziz City for Science and Technology, Saudi Arabia*.
- Al-Qadi, I.L., Xie, W. and Roberts, R. (2008). Time-frequency approach for ground penetrating radar data analysis to assess railroad ballast condition. *Research in Nondestructive Evaluation*, vol. 19, no. 4, pp. 219–237.
- Ascher, H. and Feingold, H. (1984). *Repairable systems reliability: modeling, inference, misconceptions and their causes*. M. Dekker New York.
- Backlund, F. and Hannu, J. (2002). Can we make maintenance decisions on risk analysis results? *Journal of Quality in Maintenance Engineering*, vol. 8, no. 1, pp. 77–91.
- Barrett, G.B. (2000). The coefficient of determination: Understanding r^2 and R^2 . *The Mathematics Teacher*, pp. 230–234.
- Bertossi, L., Kolahi, S. and Lakshmanan, L.V. (2013). Data cleaning and query answering with matching dependencies and matching functions. *Theory of Computing Systems*, vol. 52, no. 3, pp. 441–482.
- Börjesson, M. and Eliasson, J. (2011). On the use of ‘average delay’ as a measure of train reliability. *Transportation Research Part A: Policy and Practice*, vol. 45, no. 3, pp. 171–184.
- Chin, K.S., Wang, Y.M., Poon, G.K.K. and Yang, J.B. (2009). Failure mode and effects analysis by data envelopment analysis. *Decision Support Systems*, vol. 48, no. 1, pp. 246–256.
- Coetzee, J.L. (1997). The role of NHPP models in the practical analysis of maintenance failure data. *Reliability Engineering & System Safety*, vol. 56, no. 2, pp. 161–168.
- Coetzee, J.L. (2004). *Maintenance*. Trafford Publishing, Victoria.
- Conradie, P.D.F. (2015). *Quantifying System Reliability in Rail Transportation in an Ageing Fleet Environment*. Master’s thesis, Stellenbosch University.

- Crow, L.H. (1990). Evaluating the reliability of repairable systems. In: *Reliability and Maintainability Symposium, 1990. Proceedings., Annual*, pp. 275–279. IEEE.
- Crowder, M.J., Kimber, A., Sweeting, T. and Smith, R. (1994). *Statistical analysis of reliability data*, vol. 27. CRC Press.
- Cryer, J.D. and Kellet, N. (1986). *Time series analysis*, vol. 101. Springer.
- Elsayed, E.A. (2012). *Reliability engineering*, vol. 88. John Wiley & Sons.
- Esveld, C. (2001). *Modern Railway Track*. 2nd edn. MRT Productions, Zaltbommel.
- Evans, A.W. (2011). Fatal train accidents on Europe’s railways: 1980–2009. *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 391–401.
- Fogel, G. (2013). It is time to start having the right conversations about ISO 55000. *Uptime - for maintenance reliability and asset management professionals*, vol. 7.
- Geum, Y., Seol, H., Lee, S. and Park, Y. (2009). Application of fault tree analysis to the service process: service tree analysis approach. *Journal of Service Management*, vol. 20, no. 4, pp. 433–454.
- Gevert, T. (2007). Carajas rises to capacity challenge. *International Railway Journal*, vol. 47, no. 6.
- Golmakani, H.R. and Pouresmaeeli, M. (2014). Optimal replacement policy for condition-based maintenance with non-decreasing failure cost and costly inspection. *Journal of Quality in Maintenance Engineering*, vol. 20, no. 1, pp. 51–64.
- Greve, M., Dersch, M.S., Edwards, J.R., Barkan, C.P., Mediavilla, J. and Wilson, B.M. (2014). Analysis of the relationship between rail seat load distribution and rail seat deterioration in concrete crossties. In: *2014 Joint Rail Conference*. American Society of Mechanical Engineers.
- Grobbelaar, S. and Visser, J. (2015). Determining the cost of predictive component replacement in order to assist with maintenance decision-making. *South African Journal of Industrial Engineering*, vol. 26, no. 1, pp. 150–162.
- Guo, H., Liao, H., Pan, R., Aron, A. and Mettas, A. (2010). Pathological behaviors of Fisher confidence bounds for Weibull distribution. *Computers & Industrial Engineering*, vol. 58, no. 4, pp. 717–728.
- Harnly, J.A. (1998). Risk based prioritization of maintenance repair work. *Process Safety Progress*, vol. 17, no. 1, pp. 32–38.
- Higgins, A. (1998). Scheduling of railway track maintenance activities and crews. *Journal of the Operational Research Society*, vol. 49, no. 10, pp. 1026–1033.
- Iwnicki, S. (2006). *Handbook of railway vehicle dynamics*. CRC press.

- Jafarian, E. and Rezvani, M. (2012). Application of fuzzy fault tree analysis for evaluation of railway safety risks: an evaluation of root causes for passenger train derailment. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 226, no. 1, pp. 14–25.
- Khan, F., Haddara, M., Krishnasamy, L. *et al.* (2008). A new methodology for risk-based availability analysis. *Reliability, IEEE Transactions on*, vol. 57, no. 1, pp. 103–112.
- Kumar, S., Gupta, S., Ghodrati, B. and Kumar, U. (2010). An approach for risk assessment of rail defects. *International Journal of Reliability, Quality and Safety Engineering*, vol. 17, no. 4, pp. 291–311.
- Kyriakidis, M., Hirsch, R. and Majumdar, A. (2012). Metro railway safety: An analysis of accident precursors. *Safety science*, vol. 50, no. 7, pp. 1535–1548.
- La Band, R. and Andrews, J. (2004). Phased mission modelling using fault tree analysis. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 218, no. 2, pp. 83–91.
- Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, New York.
- Lewis, P.A. and Robinson, D.W. (1974). Testing for a monotone trend in a modulated renewal process. *Reliability and Biometry*, pp. 163 – 182.
- Lisnianski, A., Frenkel, I. and Ding, Y. (2010). *Multi-state system reliability analysis and optimization for engineers and industrial managers*. Springer Science & Business Media.
- Liu, X., Barkan, C.P. and Saat, M.R. (2011). Analysis of derailments by accident cause. *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2261, no. 1, pp. 178–185.
- Louit, D.M., Pascual, R. and Jardine, A.K. (2009). A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data. *Reliability Engineering & System Safety*, vol. 94, no. 10, pp. 1618–1628.
- Maluleke, K.J. (2013). Instilling safety culture in the passenger rail transport industry within the South African context. *Journal of Transport and Supply Chain Management*, vol. 7, no. 1, pp. 1–6.
- Markowski, A.S. and Mannan, M.S. (2008). Fuzzy risk matrix. *Journal of hazardous materials*, vol. 159, no. 1, pp. 152–157.
- Marquis, B.P., Muhlanger, M. and Jeong, D.Y. (2011). Effect of wheel/rail loads on concrete tie stresses and rail rollover. In: *ASME 2011 Rail Transportation Division Fall Technical Conference*, pp. 143–150. American Society of Mechanical Engineers.

- MIL-HDBK-189 (1981). *Reliability Growth Management*. Headquarters, U.S. Army Communications Research and Development Command, Fort Monmouth.
- Minsili, L. S., Jeremie, M. D., Tsebo Simo, G. L. and Simo, C. (2012). Preventive maintenance of railway tracks: Ballast performance anticipation in the Cameroon railway. *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 5, pp. 398–406.
- Modarres, M., Kaminskiy, M.P. and Krivtsov, V. (2009). *Reliability engineering and risk analysis: a practical guide*. CRC press.
- Mohammadzadeh, S., Ahadi, S. and Nouri, M. (2014). Stress-based fatigue reliability analysis of the rail fastening spring clip under traffic loads. *Latin American Journal of Solids and Structures*, vol. 11, no. 6, pp. 993–1011.
- Mohammadzadeh, S. and Ghahremani, S. (2012). Estimation of train derailment probability using rail profile alterations. *Structure and Infrastructure Engineering*, vol. 8, no. 11, pp. 1034–1053.
- Mohammadzadeh, S., Sangtarashha, M. and Molatefi, H. (2011). A novel method to estimate derailment probability due to track geometric irregularities using reliability techniques and advanced simulation methods. *Archive of Applied Mechanics*, vol. 81, no. 11, pp. 1621–1637.
- Morant, A., Larsson-Kråk, P.O. and Kumar, U. (2014). Data-driven model for maintenance decision support: A case study of railway signalling systems. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, pp. 1–15.
- Mundrey, J. (2010). *Railway track engineering*. Tata McGraw-Hill Education.
- Ng, M., Tummala, V.R. and Yam, R.C. (2003). A risk-based maintenance management model for toll road/tunnel operations. *Construction Management and Economics*, vol. 21, no. 5, pp. 495–510.
- Nurmikolu, A. (2012). Key aspects on the behaviour of the ballast and substructure of a modern railway track: research-based practical observations in Finland. *Journal of Zhejiang University SCIENCE A*, vol. 13, no. 11, pp. 825–835.
- Oyama, T. and Miwa, M. (2006). Mathematical modeling analyses for obtaining an optimal railway track maintenance schedule. *Japan Journal of Industrial and Applied Mathematics*, vol. 23, no. 2, pp. 207–224.
- Paté-Cornell, E. (2002). Finding and fixing systems weaknesses: Probabilistic methods and applications of engineering risk analysis. *Risk Analysis*, vol. 22, no. 2, pp. 319–334.
- Podofillini, L., Zio, E. and Vatn, J. (2006). Risk-informed optimisation of railway tracks inspection and maintenance procedures. *Reliability Engineering & System Safety*, vol. 91, no. 1, pp. 20–35.

- Prasad, C. (2012). Functional failure analysis of a Pandrol clip. *Indian Journal of Electronic and Electrical Engineering*, vol. 1, no. 1-2.
- Prescott, D. and Andrews, J. (2013). A track ballast maintenance and inspection model for a rail network. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of risk and reliability*, vol. 227, no. 3, pp. 251–266.
- Rausand, M. and Høyland, A. (2004). *System reliability theory: models, statistical methods, and applications*, vol. 396. John Wiley & Sons.
- Regulator, R.S. (2014). RSR media release.
Available at: https://www.google.co.za/webhp?sourceid=chromeinstant&rlz=1C1BLWB_enZA597ZAZ97&ion=1&espv=2&ie=UTF-8#q=PRASA+speed+restrictions+newspaper&spell=1
- Remennikov, A., Kaewunruen, S. and Ikaunieks, K. (2006). Deterioration of dynamic rail pad characteristics. University of Wollongong.
- Riera-Ledesma, J. and Salazar-González, J.J. (2007). A heuristic approach for the continuous error localization problem in data cleaning. *Computers & Operations Research*, vol. 34, no. 8, pp. 2370–2383.
- Sachs, L. (2013). *Angewandte statistik: anwendung statistischer methoden*. Springer-Verlag.
- Sadeghi, J. and Askarinejad, H. (2009). An investigation into the effects of track structural conditions on railway track geometry deviations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 223, no. 4, pp. 415–425.
- Sadeghi, J. and Askarinejad, H. (2010). Development of improved railway track degradation models. *Structure and Infrastructure Engineering*, vol. 6, no. 6, pp. 675–688.
- Sadeghi, J. and Askarinejad, H. (2011). Development of track condition assessment model based on visual inspection. *Structure and Infrastructure Engineering*, vol. 7, no. 12, pp. 895–905.
- Saldanha, P.L., de Simone, E.A. and e Melo, P.F. (2001). An application of non-homogeneous Poisson point processes to the reliability analysis of service water pumps. *Nuclear engineering and design*, vol. 210, no. 1, pp. 125–133.
- Sameni, M, K. (2012). *Railway Track Capacity: Measuring and Managing*. Ph.D. thesis, University of Southampton, Faculty of Engineering and Environment, Transportation Research Group.
- Shalev, D.M. and Tiran, J. (2007). Condition-based fault tree analysis (CBFTA): A new method for improved fault tree analysis (FTA), reliability and safety calculations. *Reliability Engineering & System Safety*, vol. 92, no. 9, pp. 1231–1241.

- Shcherbanin, Y.A. (2012). Some problems of Russia's railway infrastructure. *Studies on Russian Economic Development*, vol. 23, no. 1, pp. 37–47.
- Silvast, M., Nurmikolu, A., Wiljanen, B. and Levomaki, M. (2010). An inspection of railway ballast quality using ground penetrating radar in Finland. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 224, no. 5, pp. 345–351.
- Song, D.-P. (2009). Production and preventive maintenance control in a stochastic manufacturing system. *International Journal of Production Economics*, vol. 119, no. 1, pp. 101–111.
- Stenstrom, C., Parida, A., Galar, D. and Kumar, U. (2013). Link and effect model for performance improvement of railway infrastructure. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, pp. 392–402.
- Sun, W., Whelan, B., McBratney, A.B. and Minasny, B. (2013). An integrated framework for software to provide yield data cleaning and estimation of an opportunity index for site-specific crop management. *Precision agriculture*, vol. 14, no. 4, pp. 376–391.
- Topham, G. (2014). Network rail fines more than £50m for late trains.
Available at: <http://www.theguardian.com/business/2014/jul/07/network-rail-fined-50m-pounds-late-trains>.
- US Dept. of Transportation (2011). Effect of missing or broken fasteners on gage restraint of concrete ties. Research results - FRA office of railroad policy and development.
- Van den Broeck, J., Cunningham, S.A., Eeckels, R. and Herbst, K. (2005). Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS medicine*, vol. 2, no. 10, p. 966.
- Van der Merwe, G. and Venter, P. (2001). IM2000 infrastructure measuring car: the application of recording results. *Rail Engineering International*, vol. 30, no. 4.
- Vlok, P.J. (2012). Introduction to elementary statistical analysis of failure time data: Long term cost optimisation and residual life estimation. Class notes.
- Wang, P. and Coit, D.W. (2005). Repairable systems reliability trend tests and evaluation. In: *Proceedings of the Annual Reliability and Maintainability Symposium*, pp. 416–421. Citeseer.
- Wang, Y.M., Chin, K.S., Poon, G.K.K. and Yang, J.B. (2009). Risk evaluation in failure mode and effects analysis using fuzzy weighted geometric mean. *Expert Systems with Applications*, vol. 36, no. 2, pp. 1195–1207.
- Woo, C.S. and Park, H.S. (2014). Useful lifetime prediction of rail pads for high speed trains. *International Journal of Mechanical, Aerospace, Industrial and Mechatronics Engineering*, vol. 8, no. 4.

- Xiao, N., Huang, H.Z., Li, Y., He, L. and Jin, T. (2011). Multiple failure modes analysis and weighted risk priority number evaluation in FMEA. *Engineering Failure Analysis*, vol. 18, no. 4, pp. 1162–1170.
- Xu, K., Tang, L.C., Xie, M., Ho, S. and Zhu, M. (2002). Fuzzy assessment of FMEA for engine systems. *Reliability Engineering & System Safety*, vol. 75, no. 1, pp. 17–29.
- Xu, P., Sun, Q., Liu, R. and Wang, F. (2011). A short-range prediction model for track quality index. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 225, no. 3, pp. 277–285.
- Zaayman, L. (2011). Track condition monitoring and analysis for effective maintenance planning. *Civil Engineering: Magazine of the South African Institute of Civil Engineering*, vol. 19, no. 4, pp. 25–33.
- Zaayman, L. (2013). *Mechanisation of track work in developing countries*. Independent, South Africa.
- Zhao, J., Chan, A. and Burrow, M. (2007). Reliability analysis and maintenance decision for railway sleepers using track condition information. *Journal of the Operational Research Society*, vol. 58, no. 8, pp. 1047–1055.
- Zixian, L., Xin, N., Yiliu, L., Qinglu, S. and Yukun, W. (2011). Gastric esophageal surgery risk analysis with a fault tree and markov integrated model. *Reliability Engineering & System Safety*, vol. 96, no. 12, pp. 1591–1600.